Technical and Policy Issues Surrounding the Use of Autonomous Maneuverable Earth Observing Satellites

by

Joshua B. McConnell

M.S., Mechanical Engineering
Texas A&M University (2000)

B.S., Mechanical Engineering
Texas A&M University (1997)

Submitted to the Department of Aeronautics and Astronautics and the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Aeronautics and Astronautics
and
Master of Science in Technology and Policy

at the
Massachusetts Institute of Technology
June 2003

Author ....
Department of Aeronautics and Astronautics and
Technology and Policy Program, Engineering Systems Division
May 16, 2003

Certified by ....
Brian Williams
Professor of Aeronautics and Astronautics
Technical Thesis Supervisor

Certified by ....
Daniel Hastings
Professor of Aeronautics and Astronautics and Engineering Systems
Policy Thesis Supervisor

Accepted by..
Daniel Hastings
Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program
Chair, Committee on Graduate Studies

Accepted by..
Edward. M. Greitzer
H.N. Slater Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

Technical and Policy Issues Surrounding the Use of Autonomous Maneuverable Earth Observing Satellites

by

Joshua B. McConnell

Submitted to the Department of Aeronautics and Astronautics and the
Engineering Systems Division
on May 16, 2003 in partial fulfillment of the requirements for the Degrees of
Master of Science in Aeronautics and Astronautics and
Master of Science in Technology and Policy

Abstract

To better respond to transient Earth phenomenon that can cause loss of life or damage to economic assets (tornadoes, mudslides, flash floods, etc.), an increase in the amount and timeliness of information collected on phenomenon is needed. One method for collecting this information is by using groups of Earth observing satellites with the ability to perform autonomous orbital maneuvers and view phenomenon on demand. However, as satellites are very costly, creating a group of satellites large enough to perform this task is currently beyond the abilities of any one organization. One method of gathering a group of satellites that is large enough is by several organizations "pooling" their satellite resources together temporarily.

In order to pool autonomous maneuverable satellites, several technical and policy problems must be overcome. The technical problem addressed is how to schedule large numbers of satellites to effectively collect critical information on phenomenon, even in the face of unexpected events, such as satellite failures that can prohibit the collection of this information. The policy problem addressed is how to overcome barriers that prevent organizations from temporarily loaning their satellite resources to a pooling system.

To overcome the technical problem of effectively scheduling large numbers of satellites, an integrated planner is developed using Draper Laboratory's EPOS 1.0 optimal planner and the ALLIANCE behavioral planning algorithm. The optimal planner efficiently allocates satellite and fuel resources, while the reaction planner modifies the optimal plan if an unexpected event occurs that would decrease the group's ability to collect information.

To overcome the policy problem of assembling a large number of satellites, a public-private partnership pooling organization is proposed. As satellites are currently a highly expensive and limited resource, the willingness and ability of organizations with satellite resources to contribute part of their satellite resources is in question. Barriers identified when forming a pooling organization and ways to overcome these barriers are identified.

Through the analysis of several simulations it was found that it is possible to achieve the technical results of responding to unexpected events in a timely manner without a substantial increase in fuel usage. Through a policy analysis it was determined that the liability issues associated with satellite pooling and organizational cultural inertia are the primary barriers inhibiting organizations from participating in a pool, but that these are possible to overcome, as there are examples where similar cross organizational relationships have succeeded with great effort. This thesis finds that the critical barriers that must be resolved before creating a group of autonomous maneuverable Earth observing satellites are not technical in nature, but are legal and cultural changes in organizations.

Thesis Supervisor (Technical): Brian Williams
Title: Professor of Aeronautics and Astronautics

Thesis Supervisor (Policy): Daniel Hastings
Title: Professor of Aeronautics and Astronautics and Engineering Systems

Acknowledgements

The author would like to thank the Decision Systems Group at Draper Laboratory for providing the project and funding that made this work possible. Thanks go to Mark Abramson, who served as my supervisor of the technical work completed at Draper Laboratory and reported in this thesis. Special thanks go to Dave Carter, who was patient enough to sit and explain the finer points of orbital mechanics and the inner workings of various computer code for hours on end on more than one occasion.

Special thanks go to Rob, Andy, Trent and Tony, the Draper Fellows who made the time in and out of the lab especially enjoyable. Thanks to my roommates Rob, Andy, Jason and Laura for putting up with my weird hours during the writing of the thesis and constant state of exasperation. A very special gratitude owed to Susan, for re-opening my eyes to the fact that there is so much more to life than research and theses. And finally, many thanks to my family; Franci, Dennis, Shannon, Rachel, Ethan, Rick and Maxine, for all the love and support that has been given, not just while working on this thesis, but for my entire life.

The technical content of this report in its entirety was prepared at The Charles Stark Draper Laboratory, Inc., under Contract #NAS3-00163, sponsored by NASA's Earth Science Technology Office. The technical content alone can be found as Draper Laboratory Report Number R-2936.

Publication of this report does not constitute approval by Draper Laboratory or the sponsoring agency of the findings or conclusions herein. It is published for the exchange and stimulation of ideas.

Joshua B. McConnell

May 16, 2003

Contents

List of Figures

List of Tables

Chapter 1

Introduction

From the Norse Thunder gods to the modern day weatherman, people have attempted to explain and understand the powerful and sometimes destructive forces of the natural events and phenomenon that surround all humans. With varying degrees of success humans try their best to shield themselves, others and personal belongings from the ravages of the natural world.

The risks posed to people and economic assets stemming from Earth-based phenomenon, which are defined in this thesis as either natural disasters or disasters stemming from the malfunction of human systems, are substantial. Currently, naturally occurring phenomenon like hurricanes, mudslides and tornadoes and human caused phenomenon like oil spills, forest fires and pollution cause a great amount of deaths, injuries and destruction around the globe annually. The annual average direct economic cost due to just hurricanes, tornadoes and flood damage in the United States is estimated at over $11B (measured in 1999 dollars) [33] while close to 3000 people in just the United States are either killed or injured annually due to various natural disasters [31].

Humans have long tried to understand these various phenomenon so that this knowledge could then be harnessed and applied, in order that the negative effects that they have on people and property can be reduced. This acquisition of knowledge has met with various degrees of success. In the quest to increase our understanding, theories on how these phenomenon behave have been created and many observations have been made to gather information on them. Three main reasons exist that drive the need to obtain information that will allow an effective response to Earth based phenomenon:

- increase the ability to aid people that may be harmed by the phenomenon,
- increase the ability to protect property or assets that may be damaged or destroyed by the phenomenon and
- increase the understanding of Earth based phenomenon.

In order to aid our ability to gather information through observations, people have employed many tools and sensors that serve to gather data in greater amounts, detail and form than is possible using human senses alone.

One of the tools that have been employed in relatively recent years to aid in gathering new information is the Earth observing satellite. Satellites have been used successfully to increase our understanding of various types of phenomenon and to help mitigate the damage caused from them by using the satellite's unique vantage point and information gathering characteristics. It is anticipated that satellites will play an even more vital role in the future for understanding and mitigating disastrous consequences of various phenomenon. In order to do this, additional capabilities not currently available must be researched, designed, developed and made operational.

A quote from former NASA administrator Daniel Goldin serves to illustrate a vision of what capabilities are needed for the future.

> "...Thus far, we are only experimenting with long term weather, climate, and natural hazard prediction. The quest for a true predictive capability for Earth system changes *requires a flexible and progressive space system architecture that is responsive* to our needs based on our current understanding of the system as well as accommodating emerging needs in the coming decades. We need to design and establish a smart, *autonomous and flexible constellation* of Earth observing satellites which can be

reconfigured based on the contemporary scientific problems at hand. Such a constellation would exploit a combination of active and passive sensing sensors in ways that we can perhaps imagine today....." (italics added for emphasis)

This quote defines a need to better predict and respond to various types of Earth based phenomenon through the development and operation of satellite systems that will provide the ability to gather the proper types of information through observations. Capabilities to accomplish this include the ability for satellites to cooperate and gather information as a group in a coordinated manner, even in the face of unexpected failures and the occurrence of unplanned events.

The ability for groups of satellites to gather information in a coordinated manner is an essential step in meeting the challenge of increasing the amount, timeliness and type of observations that are currently "incomplete in time and space." [37]. Currently, the vast majority of Earth observing satellites perform observations either in solitary or as part of a static constellation. A static constellation is defined here as being composed of satellites that maintain their relative positions between one another and do not engage in orbital maneuvers outside of position keeping chores. All satellites are restricted to the orbit that they are launched into, with infrequent or non-existent opportunities to change the orbit. While the orbits usually provide a repeat ground track, the time before a satellite will once again view a specified target is often on the order of days. For example, the Landsat 7 satellite launched in mid-1999 had a repeat coverage time of 16 days [31], meaning that it would not see a specific target again for 16 days.

As the number of targets that are of interest increases and the amount and timeliness of the information required for each target increases, the revisit times on the order of days that are possible with static constellations are not adequate to obtain the information that is required. Some combination of enabling the satellites to visit a target on a more frequent basis or providing more satellites that can see the target is necessary, if satellites are to be tasked with the responsibility of increasing the amount of target observations.

In order to achieve either decreased revisit times or support of more satellites, new technologies must be developed and new ways of managing groups of satellites must be implemented.

Satellite Technology

To decrease target revisit times or make the deployment of greater numbers of satellites feasible, many technologies must be developed to an extent greater than they exist today. These technologies cover a broad range of fields and include miniaturization of spacecraft systems and sensors, better data networks and communications management, increased orbital control and operational planning and the ability to launch satellites with greater frequency at lower costs [27]. Improvements in these areas have been recognized as being necessary by a variety of organizations that engage in Earth observations, including the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA), which funds research into these areas, including work completed for this thesis. While technologies are utilized as they become available, it is a goal that in the 2010 to 2025 time frame a "proactive environmental prediction" system will be in place [27].

In determining which, if any, of these technologies should be implemented to increase the amount and timeliness of information collected on Earth phenomenon, more work on understanding the different alternatives in each of these technologies must be conducted. One orbital control and operational planning technology alternative, the integrated planner, is studied in this thesis. The integrated planner combines an optimal planner [1] and a reaction planner [17, 18, 19, 20, 21]. The purpose of the optimal planner is to efficiently allocate the limited satellite and fuel resources available. Efficiency is measured here in the amount of observation time that can be achieved with an amount of fuel used, with highly efficient systems obtaining large amounts of timely observations while only using small amounts of fuel. The purpose of the reaction planner is to modify the plan that is produced by the optimal planner if an unexpected event occurs that causes a decrease in the group's ability to collect information. The reaction planner will then re-task satellites in the group to regain as much of the lost observation time as is possible. The integrated planner uses the previously developed Draper Laboratory's EPOS 1.0 optimal planner [1] and the previously developed ALLIANCE behavioral planning algorithm [17, 18, 19, 20, 21]. It should be noted that the Draper EPOS 1.0 optimal planner is actually a combination of planner and scheduler, as both the tasks that need to be accomplished as well as when and what satellites will accomplish the tasks are determined. For simplicity, the EPOS 1.0 optimal planner and the integrated planner will only be referred to as a planner.

The main technical focus of this research is in examining the utility of the reaction planner in providing a modified plan to regain the information lost after the occurrence of an unexpected event. It is believed that without the capabilities provided by the reaction planner, a substantial amount of information would be lost after the occurrence of an unexpected event. Therefore, it is believed that a reaction planner is necessary to regain information lost after an unexpected event. This thesis examines whether the reaction planner can re-task the satellite group to allow a substantial amount of information to be regained in a timely manner, without using a substantially larger proportion of fuel than was originally allotted. This is studied in terms of the percent of observation time that is regained, the time elapsed between the unexpected event and the resumption of observations and the amount of fuel that is needed to provide these new observations.

Satellite Management

To increase the number of satellites that could be tasked to observe an Earth based phenomenon on demand, new ways of managing satellites must be developed. Currently, satellite command and control functions are executed by a single organization with operational responsibility for the satellites in their care. There is little coordination across satellite operators or between satellites under the same operator's control. This lack of coordination is due not only to current satellite design limitations, but organizational limitations as well. There is little history, motivation or reason for most organizations to coordinate the activities of their satellites with the activities of other organization's satellites. Part of the reason that satellite coordination is so difficult is because so many stakeholders are involved. Satellite owners include government research agencies (such as NASA), government operational agencies (such as NOAA), military organizations, intelligence organizations, international governments, universities and private industry. Often, satellite owners, operators and users are all different groups as well. Coordination amongst these varied stakeholders may be necessary if the "proactive environmental prediction" system envisioned for a 2010 – 2025 deployment will become a reality. [27]

Unless satellite technology increases so that it becomes feasible that the financial resources of one organization will be great enough to deploy the large number of dedicated satellites necessary to achieve the small target revisit times that are envisioned, satellite resources from several organizations will have to be used. One alternative for using satellites from several different organizations is to "pool" satellite resources together across organizations. The pool envisioned in this thesis enables organizations to allow their satellites to be used on a temporary basis, when the satellite is not engaged in a task of critical importance for the organization with primary satellite ownership. The organizational, political, economic and legal hurdles for creating and implementing a satellite pooling system are studied in this thesis.

The main policy focus of this research is in examining the barriers that exist that inhibit organizations from participating in a satellite pool. It is believed that significant benefits can be achieved by enabling satellite coordination through pooling, but that large barriers exist that make it costly for organizations to participate in the pool. Organizational models for creating an organization for the purposes of creating and operating a satellite pooling system capable of overcoming identified barriers is also studied. This thesis examines barriers and organizational models capable of overcoming these barriers for the creation of a pooling organization capable of creating and operating a pooling system.

Thesis Objective

This thesis studies one alternative design for creating a group of satellites capable of observing Earth based phenomenon in real time. The design includes both a technology component (the integrated planner) and a policy component (creation of a satellite pool). The objective of this thesis is to determine if the studied design is both technologically and politically feasible.

The technology and policy aspects of this problem are tightly coupled together. The choice of using a group of autonomous maneuverable satellites indicates that these satellites will be large and will have the capacity to be refueled, which is the only way to make orbital maneuvers feasible on the time scale needed. The use of groups of large satellites drives the question of where will the satellites that form the groups come from. As these satellites are very costly, it is currently beyond the resources of any organization to create a group with enough satellites to fulfill the mission of collecting information on phenomenon. This economic reality drives thinking towards different ways of creating groups of satellites, such as the satellite pooling system. However, as the pooling system constitutes a new way of operating satellites, this poses the policy problem of how to actually get organizations .to participate in a satellite pool. As coordinating so many organizations across so many different sectors may prove to be too great a challenge, this policy problem may force a drive to develop new technologies, such as microsatellites, that could address the technical problem without sparking a major policy problem as well.

Thesis Organization

This thesis deals with technical and policy problems pertaining to observing Earth based phenomenon in near-real time with autonomous maneuverable satellites. These problems are studied in the ten chapters contained in this thesis. Chapter 1 provides a high level introduction to the problem and the underlying motivation for why this problem was studied. Chapter 2 elaborates on the problems being studied and provides and overview of how the problems were attacked. Chapter 3 provides an overview of background information necessary for the study of these problems. Chapters 4-6 detail the technical research and results found for the technical aspect of this problem. Chapters 7-8 detail the policy problem and results found for the policy aspect of this problem. Chapter 9 provides a discussion of the technical and policy results in terms of the problem defined in Chapters 1 and 2. Chapter 10 provides conclusions drawn from the work of the entire thesis.

Chapter 2

Problem Definition and Research Objectives

This chapter provides an overview of the identification and definition of the technical and policy problems associated with autonomously observing various Earth based phenomena for the purpose of providing real-time information. This chapter expresses the underlying need behind this research, discusses some critical aspects of the problem in more detail, and graphically maps the driving need to the focus of this research. Addressing the technical aspect of this research, a concept trade space, built around previous NASA studies [27], is presented which displays several types of systems designed to provide observations of Earth phenomenon. From these concepts a dynamic constellation, which is a set of satellites capable of performing autonomous orbital maneuvers, is down selected for further study. Addressing the policy aspect of this research, a concept trade space is presented which displays several methods of gathering and managing a group of satellites large enough to provide observations of Earth phenomenon. From these concepts a pool system, defined as temporarily placing satellites into a common pool for use by a pooling organization, is down selected for further study. The overall research goal of determining if the proposed technical and policy solution for observing Earth phenomenon is feasible is further elaborated on. The technical portion of the research goal consists of quantifying benefits and costs derived from use of the integrated planner. The policy portion of the research goal consists of identifying and overcoming barriers that exist to pooling satellite resources. The chapter is concluded with an overview of the procedure used in this research.

2.1 Need Statement

A perceived need has been identified, which can be stated as:

There is a need to efficiently obtain information to aid in effectively responding to various Earth based phenomenon.

Explaining terms in the need statement, "efficiently obtain information" means providing the greatest amount of information possible for a set amount of fuel and "effectively responding to various Earth based phenomenon" means to task the satellite group so that it can obtain a set of observations that at least meets the threshold of the minimum amount of observations required to be useful. Part of being effective is ensuring that the number of satellites that are required to obtain the needed observations are available. It is assumed that "efficiency" and "effectiveness" will be quantitatively defined on a mission by mission basis, with the amount of observations needed, the amount of fuel willing to be expended, and the number and types of satellites required dependant on the priority of the target and the satellites that are available to view the target.

2.2 Need Analysis

Based on the above need statement, the problem that would ideally be solved is one where all the information that is desired could be collected for an arbitrary number of different types of targets. This information would then be of use so that people could either predict or respond to the phenomenon being observed. This would mean that the proper type of information could be collected, when it was needed and then parceled out in a useful form to the desired recipients in a timely manner so that it could be acted upon. This problem description would likely entail many targets, perhaps on the order of hundreds or thousands, each with a changing relative priority, which would appear, physically move, and then disappear over time. In order to view all these targets in a timely manner a large number of satellites would be required, probably on the order of hundreds. These satellites would need to be able to, among other capabilities, coordinate among themselves to determine what would be a "good" plan so that the required observations of the selected targets could be obtained.

Here the terms "required observations" and "selected targets" are defined as follows. A phenomenon is identified as being interesting because of the phenomenon's impact on human life, property or scientific value. When this interest is great enough it drives a desire to collect more information on the phenomenon and the phenomenon becomes a selected target for the satellite group to observe. To gather the minimal amount of information necessary to provide useful information about the selected target, the satellite group must perform some base set of observations, which are the required observations. It is expected that more targets and observations will be desired than are possible with the satellite resources available, so some means must be provided at prioritizing what targets should be selected and what observations gathered. It is assumed here that the manner in which targets are selected are outside the scope of this thesis and that a suitable metric for determining which and how many observation should be gathered is a given for the integrated planner. However, some discussion is given on how the target type and selection may affect the willingness of organizations to contribute satellites to the pool.

Because of the short time span in which decisions would have to be made and because of the large number of satellites and targets involved, some level of autonomy in planning is needed for the system. Autonomy is required because the amount of time and effort involved in manually creating a good plan for even one satellite is not trivial [25]. It is anticipated that as the number of satellites and targets increases, the ability to create a good plan in the short time span available will not be feasible manually. Also driving the need for an autonomous planner is the necessity for the satellite group to respond to different unexpected events, such as a satellite system failure or a new viewing opportunity. As these unexpected events may prevent some of the required observations from occurring, it is essential that the plan be repaired as quickly as possible, so that as many of the original observations can be regained. This requires a fast reaction time that could not be achieved manually.

Looking now at the need statement more in-depth, two main functions of a system capable of observing various Earth based phenomenon are identified, these being the ability to respond appropriately to the phenomenon of interest and the ability to obtain information. An overview of these two functions is presented below.

Ability to Respond to Phenomenon

The ability to respond appropriately to Earth-based phenomenon is dependent on the design and deployment of a system that is, one, capable of providing understanding of the phenomenon of interest, two, has the ability to react to phenomenon of interest as they become apparent, and three, provides the ability to predict future phenomenon. An appropriate response to an Earth-based phenomenon will be dependant on what the specific phenomenon is, but may include: monitoring the phenomenon for a set amount of time (oil spills), monitoring the phenomenon as it is first appearing (tornadoes) or providing enough information through observations to be able to predict what effect the phenomenon will have on its surroundings (predicting hurricane's landfall location). Predicting or reacting accurately to phenomenon that have destructive capabilities, like hurricanes or tornadoes, is critical to the quest of saving lives and limiting economic damage due to natural disasters. To do this, however, people must know in a timely manner that a disaster is either imminent or has a high probability of occurring, and to do this, timely information on various phenomenon is required and must be obtained, which is discussed in the following section.

Two types of phenomenon are of interest to an Earth observing satellite system, these being phenomenon that are sustained for a relatively lengthy amount of time, such as the breaking of ice shelves into the formation of ice burgs, and phenomena that have a shorter temporal existence, such as hurricanes, oil spills or tornados. Targets falling into the second class are difficult to adequately observe using the static satellites and constellations available currently. The short temporal lifespan can often mean that the time between useful observations of the target can only be on the order of minutes or hours, depending on the target, as opposed to days or weeks. Targets that exist only for a relatively short period of time are the focus of this study.

Ability to Obtain Information

One of the critical methods in which the above function of responding to Earth based phenomenon is accomplished is through better understanding and observations of the phenomenon. These phenomenon can be better understood through study, combining various observations and modeling techniques. In gathering information through observations some system will need to be developed that has the ability to identify what observations and information need to be gathered, has the ability to gather this information and then possesses the ability to appropriately disseminate the gathered information to various sources. Looking at these sub-functions in greater detail:

Identify Information

The ability to identify what information should be gathered is critical. This is needed to effectively focus the proper resources to the areas of importance. It is expected that there will always be more phenomenon of interest than the resources available to observe them all. This necessitates that some method should exist that can select the critical phenomenon that need observing, most likely as a sub-set of a prioritized list of desired targets. The types of information that are identified to be collected will also impact different organization's willingness to take part in any pooling system, either as a general rule or during specific times.

Gather Information

Once the proper phenomena have been identified, it is critical that the required observations be conducted. The observations that are needed fulfill a wide range of requirements. Comparing the types of observations that are needed with what today's systems can currently provide, a proposed system should be able to provide, one, more observations of phenomena, two, observations that are more timely in relation to when the phenomenon is occurring, three, additional types of information that should be gathered, four, the capacity to provide coordinated observations, and five, observations over a longer time horizon.

It is anticipated that any system capable of providing these five functions will be highly complex. As with any complex system, failures should be expected to occur and a means should be provided for the system to still maintain its functionality in the face of these failures. The technical research in this thesis addresses a method of making one potential system more robust so that the function of gathering information can be accomplished. Assembling and operating such a complex system will also be highly challenging. The policy research in this thesis proposes one alternative for assembling such a system and identifies challenges and potential solutions so that assembly can be accomplished. Both are explained in greater detail later in the thesis.

Disseminate Information

Even gathering the proper information is not useful unless that information can then be disseminated in a timely manner to the proper sources in a form that is useful. The extremely large amounts of information that will be gathered will require novel new methods of storing, sorting, transferring and interpreting the information. Who the information will be disseminated to will also affect the willingness of organizations to participate in a satellite pool.

2.3 Function Structure

The function structure presented below in Fig. 2.1 displays the functions required to achieve the need statement presented above. The functions relevant to this study are outlined in bold and are functionally decomposed to a lower level than other functions. The lower level function that pertains directly to the technical aspect of this study, *provide operational support to the observation platforms during the occurrence of unexpected events*, is highlighted in the figure and explained below. The lower level function that pertains directly to the policy aspect of this study, *obtain use of platforms*, is also highlighted in the figure and explained below. The function structure tree is decomposed from the derived need statement, presented above, and displays the relationship between these lower level functions and the overall need.

The two basic functions that an Earth observing system should supply are: provide a means to obtain information and provide a means to effectively respond to Earth based phenomenon. To achieve the function of obtaining information, one of the primary functions that must be completed is to support operations that enable the appropriate platforms to travel to the target. On the technical side, operational support can include many functions, such as planning which satellites will view what targets and when, all aspects of attitude maneuvering, sensor tasking, etc. On the policy side, operational support includes gaining the ability to use a particular satellite from the appropriate organization. The aspects of operational support that are the focus of this thesis are: the provision of planning functions and the provision of a framework to gain permission to use the satellites. Planning functions must be provided both during normal operations and during the occurrence of an unexpected event, such as a platform failure or change in target availability. Providing a means for supporting platform operations during the course of unexpected events, specifically, the creation of a new plan that will re-task satellites, is the technical focus of this thesis. Gaining the ability to use the satellites is a necessary first step that must be accomplished before the satellites can be tasked to view the target. Providing a means for obtaining permission to use the needed satellites from the appropriate organization is the policy focus of this thesis.

Currently, planning function support is provided for through the extensive involvement of ground support personnel. In the advent of a platform failure or change in target parameters, ground support must determine what the problem is, formulate a solution and uplink the solution to the appropriate platform. This process is difficult and lengthy and may result in opportunities for observing the target being lost. It is desired that this function be provided for with a reduced or eliminated human ground support presence. This has been worked on in the past at the individual platform level during the NASA

Deep Space 1 mission, which was launched in late 1998 [32]. The technical portion of this thesis looks at a method for providing autonomous operational support during unexpected events for groups of platforms.

Currently, besides data sharing, there is little or no ability or need to share satellite resources between organizations. Organizations with a need for satellite data usually either own their own satellites, or purchase/use information collected from other satellites. To collect information on the number of targets of interest will require the coordination of more satellites than any one organization currently has the resources to manage. The policy portion of this thesis looks at a method for creating a constellation of satellites large enough to observe transient earth phenomenon.

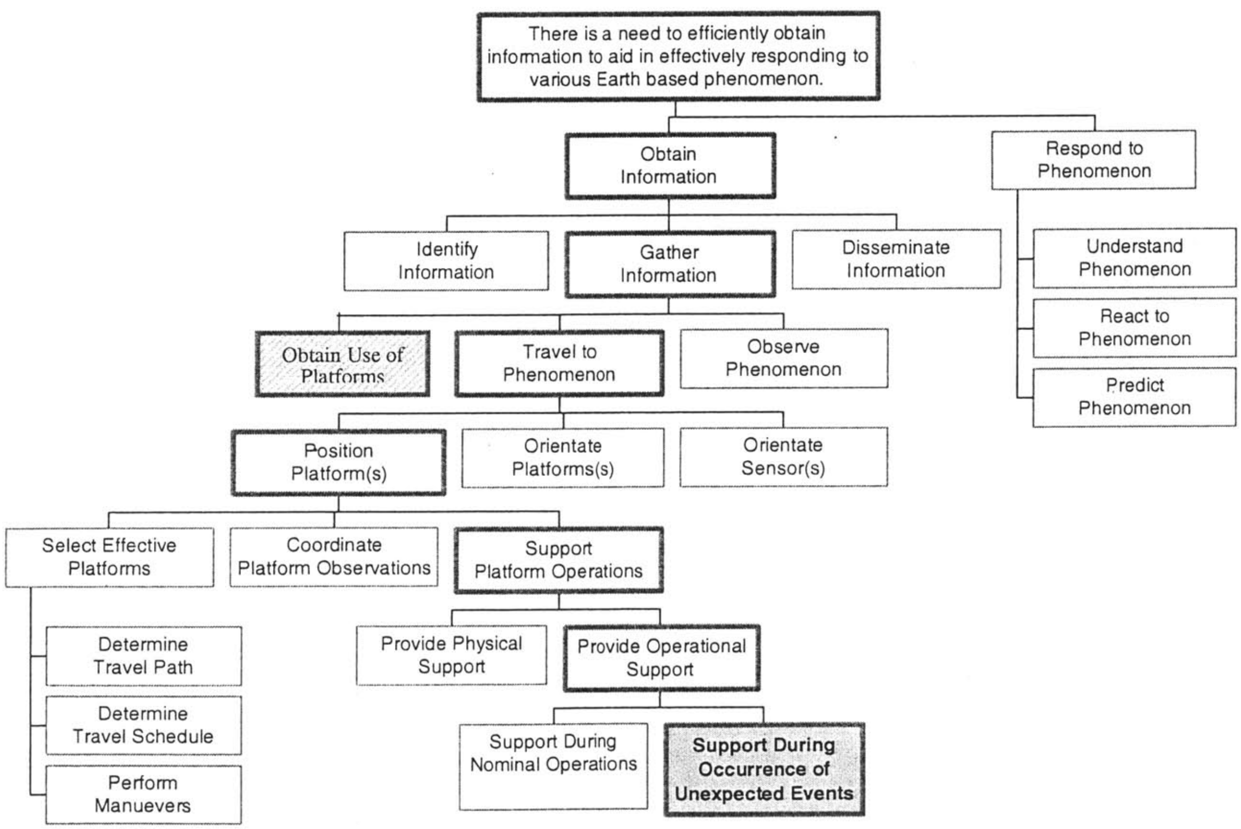


Fig. 2.1. Function structure relating need statement to research objective. Low level technical function shaded in gray, low level policy function cross-hatched.

2.4 Conceptual Designs

Two trade spaces were identified and complied. The first trade space covers technical alternatives for coordinating satellites during the occurrence of an unexpected event. The second trade space covers policy alternatives for creating a satellite group that contains enough satellites to obtain the desired information. Each trade space is discussed below in turn.

Coordination Trade Space

A trade space was identified and complied for choosing an operational support system for observation platforms during the occurrence of an unexpected event. The trade space contains information concerning possible methods of coordinating platforms and sensors. This trade space is organized by the major categories of platform coordination. Additional subdivisions expressing planning parameters and maneuvers are presented in later columns. The method chosen for study in this thesis to address the problem of providing operational support to platforms during unexpected events was a decentralized, real-time behavioral planning algorithm that was applied to a dynamic satellite constellation. This can be identified in the trade space as:

SINGLE SAT.	MANEUVERS	static	dynamic				
	SENSORS	single	multiple				
CONSTELLATIONS	INDEPENDENT OBSERVATIONS	*Maneuvers*	*Static*				
			Dynamic	Central Planning	preplan only	real time planning	
				Decentralized Planning	preplan only	real time planning	
		Sensors	*Homogenous*	single sensor / sat.			
			Heterogeneous	single sensor / sat.	multiple sensors / sat.		
	COORDINATED OBSERVATIONS	*Global Constellations*	*Maneuvers*	Static			
				Dynamic	Central Planning	preplan only	real time planning
					Decentralized Planning	preplan only	real time planning
			Sensors	Homogenous	single sensor / sat.		
				Heterogeneous	single sensor / sat.	multiple sensors / sat.	
		Virtual platform	*Central Planning*	preplan only	real time planning		
			Decent. Planning	preplan only	real time planning		
		Formation Flying					
SENSOR WEBS	SATELLITES						
	UAVs						
	GRD. STATIONS						

Fig 2.2. Platform coordination trade space.

CONSTELLATIONS → COORDINATED OBSERVATIONS → GLOBAL CONSTELLATIONS → MANEUVERS → DYNAMIC → DECENTRALIZED PLANNING → REAL TIME PLANNING

and

CONSTELLATIONS → COORDINATED OBSERVATIONS → GLOBAL CONSTELLATIONS → SENSORS → HETEROGENEOUS → SINGLE SENSOR PER SATELLITE.

The choice of these paths was driven by the desire to use the optimal planning tools available in the EPOS 1.0 optimal planner and the ALLIANCE algorithms. The selection of the types of systems represented by these paths is not presented as the best choice for use in the Earth observing satellite problem, but rather as one alternative. It is recommended that several systems be studied before determining which is the most applicable to this problem.

There are several reasons why autonomous dynamic satellite groups may be an ideal choice for observing Earth phenomenon that exist for a short time span. First, the number of opportunities that any satellite can observe a given target is limited by both the satellite sensor design and also the number of times the satellite will travel near the target. At most, satellite orbits can pass over the target twice a day, when the satellite plane and the target intersect, as shown below in Fig. 2.3. The ability to point sensors at targets or equip satellites with sensors possessing large footprints may enable a satellite to observe a target on more than these two opportunities. However, the satellite is captive to the orbit that it is placed in and may not be able to see the target at all or as often unless it can change its orbit. Because of this, satellites performing orbital maneuvers are necessary to increase the number of opportunities available to view the target.

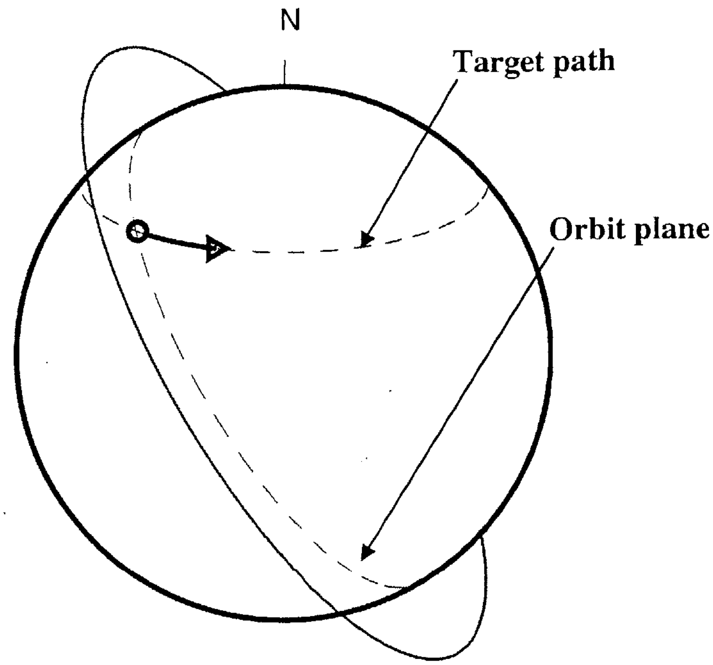


Fig. 2.3. Intersection of orbit plane and target path [43].

Second, a group of satellites will be needed to provide the amount of information that is envisioned for any one target, let alone for multiple targets. Third, as the targets are of interest for such a short period of time, it is imperative that timely information be provided, which can be accomplished by changing the orbits of satellites so that they fly over the targets sooner than they would if the orbits had remained unchanged. While prohibitively expensive with today's technology, the ability to refuel satellites while on orbit is currently be actively researched by programs such as DARPA's (Defense Advanced Research Projects Agency) Orbital Express [42]. Fourth, as the number of targets and satellites increases and the time between identification of a target and when the observations are needed decreases, the satellites should have a degree of autonomy that will allow the group to actively respond to the changing observation requirements posed by the targets. This will be required because the time available to determine what satellite should observe which target and when will be too difficult for humans to perform manually.

When determining what satellite should observe which target and when, some sort of plan should be constructed that efficiently allocates the satellite and fuel resources among the various targets of interest. While there are several planning techniques that could perform this function, an integrated planner consisting of optimal planning and reaction planning techniques was studied for this thesis. There are several reasons why a combination of optimal and reaction planning techniques may be an ideal choice for effectively managing the satellite group. First, the number of satellites and the amount of fuel available are envisioned to always be scarce compared to the possible number of targets of interest. Therefore, some means should be provided that will efficiently and effectively allocate these resources so that the maximum benefit can be obtained through their use, which is what the optimal planner attempts to accomplish. Second, because the targets identified during any mission are likely to change or the satellite group may experience various technical difficulties some means of robustness should be provided so that the mission can still be effectively accomplished. The reaction planner's decentralized approach towards providing group robustness helps ensure that the mission can still be accomplished even in the face of various unexpected events occurring throughout the mission.

Creation of Satellite Group Trade Space

A trade space was identified and complied for choosing how the group of Earth observing satellites would be formed. The trade space contains information concerning possible types of ownership of the satellites within the group[1], types of operational control of the satellite group[2], types of satellites used in the group[3], and type of satellite group

[1] All the satellite are owned by one organization, the satellites are owned by several organizations, or the satellites are owned by several organizations, but one organization owns most of the satellites within the group.

[2] Satellites controlled by a third party with no satellites in the group, satellites controlled by the primary organization or satellites controlled by a coalition of primary organizations.

[3] Large and maneuverable satellites, large and non-maneuverable satellites, or micro-satellites.

operations[4]. Each of the first columns represent one of these areas. The method chosen for study in this thesis to address the problem of creating a group of dynamic satellites large enough to provide complete coverage of Earth based phenomenon was a pooling system. The pool contains a group of satellites temporarily donated from several organizations that retain primary ownership of the satellites (primary organizations). While the satellites are temporarily in the pool, they are controlled by the pooling organization and are tasked to help provide required information on Earth based phenomenon. This can be identified in the trade space as:

Satellites from One Organization	**Large Non-maneuverable Satellites**	Static Operations, Complete Coverage	Integration of Data from Existing Operations
	Large Maneuverable Satellites	Dynamic Constellations, Complete Coverage	Dynamic Constellations, Partial Coverage	Dynamic Constellation, Integrated with Data from Existing Operations
	Micro-Satellites	Static Operations, Complete Coverage	Static Constellation, Partial Coverage	Static Constellation, Integrated with Data from Existing Operations

Satellites Evenly Distributed over Multiple Organizations	Third Party Control	***Large Non-maneuverable Satellites***	Static Operations, Complete Coverage	Integration of Data from Existing Operations	
		Large Maneuverable Satellites	Dynamic Constellations, Complete Coverage	Dynamic Constellations, Partial Coverage	Dynamic Constellation, Integrated with Data from Existing Operations
		Micro-Satellites	Static Operations, Complete Coverage	Static Constellation, Partial Coverage	Static Constellation, Integrated with Data from Existing Operations
	Primary Organization Control	***Large Non-maneuverable Satellites***	Static Operations, Complete Coverage	Integration of Data from Existing Operations	
		Large Maneuverable Satellites	Dynamic Constellations, Complete Coverage	Dynamic Constellations, Partial Coverage	Dynamic Constellation, Integrated with Data from Existing Operations
		Micro-Satellites	Static Operations, Complete Coverage	Static Constellation, Partial Coverage	Static Constellation, Integrated with Data from Existing Operations
	Control from Coalition of Primary Organizations	***Large Non-maneuverable Satellites***	Static Operations, Complete Coverage	Integration of Data from Existing Operations	
		Large Maneuverable Satellites	Dynamic Constellations, Complete Coverage	Dynamic Constellations, Partial Coverage	Dynamic Constellation, Integrated with Data from Existing Operations
		Micro-Satellites	Static Operations, Complete Coverage	Static Constellation, Partial Coverage	Static Constellation, Integrated with Data from Existing Operations

Satellites Unevenly Distributed over Multiple Organizations	Third Party Control	***Large Non-maneuverable Satellites***	Static Operations, Complete Coverage	Integration of Data from Existing Operations	
		Large Maneuverable Satellites	Dynamic Constellations, Complete Coverage	Dynamic Constellations, Partial Coverage	Dynamic Constellation, Integrated with Data from Existing Operations
		Micro-Satellites	Static Operations, Complete Coverage	Static Constellation, Partial Coverage	Static Constellation, Integrated with Data from Existing Operations
	Primary Organization Control	***Large Non-maneuverable Satellites***	Static Operations, Complete Coverage	Integration of Data from Existing Operations	
		Large Maneuverable Satellites	Dynamic Constellations, Complete Coverage	Dynamic Constellations, Partial Coverage	Dynamic Constellation, Integrated with Data from Existing Operations
		Micro-Satellites	Static Operations, Complete Coverage	Static Constellation, Partial Coverage	Static Constellation, Integrated with Data from Existing Operations
	Control from Coalition of Primary Organizations	***Large Non-maneuverable Satellites***	Static Operations, Complete Coverage	Integration of Data from Existing Operations	
		Large Maneuverable Satellites	Dynamic Constellations, Complete Coverage	Dynamic Constellations, Partial Coverage	Dynamic Constellation, Integrated with Data from Existing Operations
		Micro-Satellites	Static Operations, Complete Coverage	Static Constellation, Partial Coverage	Static Constellation, Integrated with Data from Existing Operations

Fig 2.4. Creation of satellite group trade space.

SATELLITES EVENLY DISTRIBUTED OVER MULTIPLE ORGANIZATIONS ➔ THIRD PARTY CONTROL ➔ LARGE, MANEUVERABLE SATELLITES ➔ DYNAMIC CONSTELLATIONS, COMPLETE COVERAGE

[4] Dynamic satellite groups providing complete coverage, static satellite groups providing complete coverage, data integration of existing satellites, small group of dynamic satellites providing partial coverage, small group of dynamic satellites providing partial coverage and integrated with data provided by existing satellites (use of dynamic satellites to "fill in the gaps" of current coverage).

The pooling system as a potential means of creating a group of satellites large enough to provide complete coverage was chosen primarily based on reasons of cost. It was assumed that to assemble a group of satellites large enough to observe the number of targets desired would be too costly for any one organization.[5] But, because many organizations currently have satellites already in orbit, or will be placing additional satellites in orbit, the cost to temporarily "borrow" these satellites to create a short lived group is less costly than procuring the number of needed satellites.

2.5 Research Goal

The goal of this research is to determine if groups of autonomous, maneuverable satellites formed through the pooling of various organization's satellite resources is a viable means of observing Earth based phenomenon. This goal is subdivided into two sub-goals; first, aiding in the future design of dynamic satellite groups, and two, identifying methods to create large groups of satellites from existing resources. Each is detailed below.

Technical Research Goal

Dynamic satellite groups utilizing a combination of optimal and reaction planning techniques to create and execute plans to gather observations may be a viable choice of a system capable of providing observations of various Earth based phenomenon. In order to determine if this combination of optimal and reaction planning is applicable to dynamic satellite groups, a greater understanding must be obtained of how these two types of systems would work together and what, if any, benefits could be obtained through their use. The goal of the technical research is to aid in the future design of a dynamic satellite group capable of providing needed observations of Earth based phenomenon, by beginning to identify and quantify benefits gained through

[5] A note here: the creation of one large group of satellites may be much less costly in the future, especially if micro- or nano-satellites become feasible. If this occurs, hundreds or thousands of these small satellites could be placed into orbit, reducing the need for dynamic satellite groups and coordination between organizations require for the pooling system.

implementing an integrated planner consisting of both optimal and reaction planning techniques.

Policy Research Goal

Dynamic satellite groups composed of satellites pooled from various organization's satellite resources may be a viable choice for providing observations of various Earth based phenomenon. In order to determine if satellite pooling is a viable means of creating groups of satellites, potential barriers to pooling and ways to overcome these barriers must be investigated. The goal of the policy research is to aid in the creation of future dynamic satellite groups capable of providing needed observations of Earth based phenomenon, by beginning to identify and understand barriers toward multi-organizational satellite pooling, and ways in which these barriers can be overcome.

2.6 Research Objective

A greater understanding of the challenges facing the design of a dynamic satellite group formed from pooled satellites is needed. The technical and policy research objectives to facilitate this understanding are described below.

Technical Research Objective

In order to implement a successful design of a system capable of observing Earth based phenomenon a greater understanding of the benefits that are obtained through implementing various planning techniques is required. The technical research objective proposed for this study is to identify and quantify the benefits obtained with using a system that utilizes an integrated optimal and reaction based planning system. The three sub-objectives proposed to accomplish this are as follows: 1) Study what effect optimal planning has on a group of satellites. 2) Modify and apply a reaction planning algorithm to investigate what effects implementing reaction based planning, with and without learning, has on a set of agents that possess some functional similarity with a dynamic satellite group. 3) Modifying, integrating and applying an integrated optimal and reaction based planning system to determine what effects this coupling of two types of planning systems will have on a group of Earth observing satellites. This will use the Draper EPOS 1.0 optimal planner previously developed [1] along with a previously developed reaction planner based on the ALLIANCE behavioral planning algorithms [17,18,20,21].

Policy Research Objective

In order to successfully pool satellites from multiple organizations, a greater understanding of the barriers that exist that would make pooling more difficult is required. The policy research objective for this study is to identify a method for overcoming barriers that exist towards pooling. The two sub-objectives to accomplish this are as follows: 1) Identify barriers that exist for creating a pooling system. 2) Identify methods to overcome barriers.

2.7 System Description

To study the benefits of optimal and reaction planning techniques, and the methods to overcome barriers to pooling, several simplifying assumptions were made. These assumptions include simplified representations of the: satellite system's design, operation, Earth based targets, potential unexpected event occurrence and operation of pooling system. The following is a brief overview of how these were developed. More detail can be found in subsequent chapters.

Dynamic Group Design

A dynamic satellite group was created by forming a constellation of satellites, where each satellite was given the properties of an Earth observing satellite currently in use. A Walker constellation, consisting of 6 planes with 4 satellites per plane, was the initial configuration that the satellites were placed at the start of the mission. The satellite group can be considered dynamic because once the mission begins the satellites were allowed to perform orbital maneuvers to increase their observation time of selected targets. Each satellite carries one of two types of fixed, nadir pointing sensors onboard. The two sensor types are used for gathering different types of information but were designed to possess the same field of view (FOV).

The satellite used primarily for this study was based on NASA's SeaStar satellite, a low earth orbit (LEO) satellite currently being used to gather information on oceanic induced environmental changes.

System Operation

The dynamic satellite group is allowed to perform limited orbital maneuvers autonomously. The purpose of the orbital maneuvers is to allow the satellites to perform burns that will place the satellite in a new orbit that provides increased coverage of the selected target. The type of orbital maneuver is limited to only in-plane phasing burns that are constrained to occur only at the equator, to simplify modeling the orbital mechanics needed for this study. The size of the phasing burn and when the burn occurs is determined by the optimal planner and is constrained by the goal of attempting to maximize the amount of observations that are possible while staying below a fuel usage threshold constraint. All burns are further constrained to place the satellites in orbits that produce a repeat ground track, meaning that once the satellite is over a target, it will observe the target repeatedly without having to expend any additional fuel.

Planning was conducted in two different phases. The first occurred before the mission began and utilized the optimal planner to create a plan for some of the satellites in the group to observe the chosen target. Once the mission began, these selected satellites implemented the plan. If an unexpected event occurred in the middle of the mission the reaction planner, which is used in the second phase of planning, was activated. The reaction planner would select and re-task available satellites to help fill in the gap in observation time left uncompleted because of the unexpected event.

Earth Based Targets

Different temporal phenomenon located around the globe were identified as potential targets of interest. One of these phenomenon, an Atlantic hurricane off the eastern coast of Florida, was primarily used for this study.

Unexpected Events

Unexpected events that were studied were of two types. First, problems occurring with the satellites that would disable or delay the ability to observe the target or, second, changes with the environment near the target that would prevent the satellite from observing the target, such as cloud coverage preventing visual sensors from seeing a target on the ground. A third type of unexpected event identified, the appearance of new targets or disappearance of identified targets was not studied in this thesis.

Satellite Pooling

Creating a group of dynamic satellites that can respond to transient Earth based phenomenon through pooling will require that a wide variety of satellites be available on short notice. As there will be little or no opportunity to obtain permission to use a particular satellite in real-time, it is assumed that organizations agreeing to pool their satellites would also provide a complete and current schedule of which of their satellites are available for use and when they are available for use. It is assumed that if pooling is to be viable, that organization's must continue to meet the objectives that they have been specifically tasked to accomplish (meaning that they cannot donate their satellites too often) and that there would be enough participating organizations such that the demands on any one would not be overwhelming. Further, it is assumed that an on orbit re-fueling capacity would exist that was economical enough to allow the satellite to be borrowed and returned to the original organization without a serious degradation of performance.

2.8 Research Procedure

The overall procedure that was employed for this study was, one, the development of a simulation that provides estimates on the benefits and costs derived from using groups of satellites to observe various Earth based targets, and, two, the identification of barriers to the creation of a pooling system and methods to overcome these barriers. For the technical portion of this thesis, an integrated planner was created that made use of both a previously developed optimal planning model [1] and a reaction planning system based on a previously developed behavioral planning algorithm [17,18,20,21]. Also utilized were orbital mechanics and sensor footprint models, both of which were previously developed [1]. This integrated planning simulation was also based on studying optimal planning techniques applied to single satellites and groups of satellites and developing, applying and studying a reaction based planning system applied to agents with some functional similarities to an Earth observing satellite system. For the policy portion of this thesis, barriers towards the formation of a pooling system were identified and analyzed using frameworks in political economy, economics, organizational behavior, law and politics. The creation of a pooling organization capable of creating and operating the pooling system was then considered. The formation of the pooling organization along different organizational models was considered. A public-private partnership organizational model was selected as being the best organizational model for the pooling organization and was studied in further detail.

Previous Work in Optimal Planning

The study conducted for the technical portion of this thesis is partly based on work previously completed with optimal planning for single satellites and groups of satellites with EPOS 1.0, developed at Draper Laboratory [1]. EPOS 1.0 makes use of a combination of autonomous optimal planning algorithms and user decisions to create optimal maneuver and observation plans for each satellite. The goal is to maximize the amount of time that a satellite can observe a target while staying below a threshold on fuel usage. The actual observations achieved and the amount of fuel used is based on finding the largest change in observation time to change in fuel usage. This is determined as the maximum slope in the observation as a function of fuel usage curve.

Reaction Planner

The reaction planner is a combination of a behavioral planner and a model based predictor. The behavioral planner is based on algorithms previously created for a behavioral planner, called ALLIANCE [17,18,20,21], with some modifications introduced for application to the Earth observation problem. The model based predictor is used in conjunction with the behavioral planner. The model based predictor is used to help determine the satellite best suited for viewing a given target. The results of which are given to the behavioral planner to weight all calculations.

The reaction planner was created in MATLAB and was first applied to a problem that was functionally similar to the Earth observation problem but easier to implement. The algorithms and parameters were studied and modifications were made that attempt to increase the applicability of the reaction planner to the Earth observation problem. Output from the reaction planner is used to determine what satellite should be re-tasked to replace a satellite affected by an unexpected event.

Integrated Planner

The integrated planner utilizes both the optimal planner developed for EPOS 1.0 and the reaction planner. The optimal planner was used before the mission begins in order to form an optimal plan for each satellite in the group to follow when observing a target. The reaction planner was used when an unexpected event causes one or more satellites to lose the ability perform its task of observing the target.

The integrated planner was created primarily in MATLAB, with the optimal planning algorithms being in C code. A graphical user interface is provided to enter data by the user and results are presented in graphical and numeric formats. The algorithms were applied to a set of satellites that were homogenous in terms of their orbits and sensor field of views, but heterogeneous in terms of their sensor types. The ability of the satellites to observe a target following the occurrence of an unexpected event was studied. The amount of observation time that could be recovered, the time after the occurrence of an unexpected event that new observations would begin and the fuel spent to obtain these observations were all of interest.

Barriers to Forming a Polling System

The creation of a group of dynamic satellites is likely to present policy implications beyond the technical problems inherent in the group's creation. As it is anticipated that there will be significant barriers in integrating satellites from several resources, the first step in creating a viable pooling system is identifying these barriers so that they can be overcome. The barriers that were identified fell into five general categories. These were; political economy, economic, organizational, legal and political. Critical barriers in each of these areas were identified and analyzed using techniques relevant to each field.

Creation of a Pooling Organization to Overcome Barriers of Opposition

A pooling organization capable of overcoming the identified barriers of opposition that inhibit pooling system creation and operation is a necessity to make the pooling system a reality. Different organizational models that the pooling organization could be modeled on were identified and evaluated. These organizational models were; government research agencies, government operational agencies, coalitions of existing government agencies, formation of a new government agency, academic institutions, not-for-profit institutions, private companies, and public-private partnerships. Each organizational model was studied and evaluated for applicability to the satellite pooling system problem. A public-private partnership was chosen as the most applicable for use as a pooling organization, and was studied in more detail.

Chapter 3

Background and Review of Previous Work

Only recently has there been an interest in groups of satellites cooperatively working together to provide in-depth and real-time observations of various types of Earth based phenomenon. Previous experience has focused mainly on scheduling and planning problems for individual satellites. A brief overview of some of these methods is provided in this chapter. There has been very little experience to date in designing and operating any group of cooperating, autonomous agents to accomplish a defined mission. Some research has been accomplished that attempts to solve this problem from a variety of different aspects. Two such planning techniques that may be applicable to satellite systems are optimal planning and reaction planning: One, an overview of optimal planning techniques applied by Draper Laboratory to the problem of Earth observing satellites and two, a behavioral planning and learning algorithm that may be of use to the Earth observing problem. The optimal planner is designed to efficiently allocate the limited satellite and fuel resources available and the reaction planner is designed to modify the plan produced by the optimal planner if the occurrence of an unexpected event causes a decrease in the group's ability to collect information.

3.1 Overview of Single Satellite Planning and Scheduling Techniques

Previous experience with planning and scheduling for satellite has focused primarily on single satellite applications. This is because, historically, most satellite systems have been designed to operate independently, without the need to coordinate actions with groups of satellites. This section provides an overview of some of the methods that have been developed for single satellite planning and scheduling.

Single Satellite Planning Techniques

Planning is a technique that is used to determine what specific tasks should be accomplished during the course of the mission. Scheduling, discussed in the next section, is the lower level decomposition of the plan, where the resources and specific times that agents will perform the tasks are determined. The integrated planner developed for this thesis is in reality a combination of planner and scheduler as both tasks are determined and what satellites will view the tasks when is also determined.

Relatively few planners for space systems have been developed in the past. A couple of examples of planners that have been used or are currently under development are the HSTS and ASPEN planners. The HSTS planner was an early resource driven approach to planning for the Hubble Space Telescope. The HSTS planner created an optimal schedule that responded to requests for observations. Each observation consisted of tasks that initialized the telescope for viewing, the actual observation and any "clean-up" tasks that must be accomplished. The HSTS planner then coordinated all the tasks for between each observation to create an overall plan for using the telescope to achieve a series of observations. The key contribution from the HSTS planner is the integration of the planning and scheduling functions into one process [13].

Another planning process that builds on the HSTS integrated model is ASPEN. Like HSTS, ASPEN integrates both planning and scheduling functions to translate high-level mission goals into low level commands to the agent. ASPEN also includes and iterative repair model that continuously attempts to improve the quality of the schedule that is produced. This repair model is also used to modify the schedule that is produced so that new tasks can be incorporated into the schedule as they appear or unnecessary tasks can be removed from the schedule as they become unnecessary. ASPEN is designed to continuously repair the schedule that is produced throughout the mission and is thus strongly driven by temporal constraints. Temporal constraints refer to appropriate resource utilization over time. For example, scheduling tasks that will draw a large amount of energy sequentially before batteries could be recharged is not allowed.

Single Satellite Scheduling Techniques

Scheduling is a technique that determines what specific times an agent will perform specific actions and what resources will be allocated to achieve these tasks. This ordering of actions and assignment of times is important for several reasons. First, scheduling ensures that there are no conflicts between required actions that must be performed. For example, if an observation platform, such as the Hubble Space Telescope, needed to perform an attitude maneuver to properly position the satellite to obtain an observation, scheduling would ensure that the attitude maneuver was performed before the observation was performed, not after or during the observation. Second, a schedule can be developed with the goal of optimizing the ordering of the tasks that must be completed. For example, the order of observations that the Advanced X-Ray Astrophysics – Imagining satellite performs is based on a schedule that attempts to optimize some science goal, such as maximizing time spent viewing the science targets or minimizing the amount of fuel consumed during attitude maneuvers to point the craft at different targets [25]. Third, flexible schedules can be developed. These schedules try and seamlessly integrate new tasks into the existing schedule as the new tasks appear. A flexible schedule integrates the new tasks without seriously disrupting the existing schedule.

Some basic types of scheduling that have been used in the past for single satellite applications are: manual scheduling, envelope scheduling, heuristic scheduling, and optimized scheduling [25]. A short description of each follows.

Manual scheduling was the first type of scheduling done for satellites. As the name implies, people created the schedule manually. Computer programs were developed to aid in the process but were limited to checking the schedule produced to verify no constraints were violated [25].

Envelope scheduling is a simple method for allocating satellite resources among several users. In the envelope scheduling method, satellite resources are divided into blocks, or envelopes, that are then divided amongst the users. Each user can then develop a schedule for their own envelope. The schedule produced is subject to resource constraints and system constraints. The end result is a satellite level schedule that is nearly feasible, with any conflicts resolved at the mission control level. The benefit of envelope scheduling is that since each user develops a schedule, little effort is expended at the satellite level in developing a schedule. The downside of using envelope scheduling is that is a user does not use resources that have been allocated to their envelope those resources are wasted [25].

A method similar to envelope scheduling is course graining. While envelope scheduling divides satellite resources into blocks, course graining divides satellite time into blocks. This is used when multiple users need access to a single system on the satellite and must take turns using the resource sequentially [25].

Heuristic scheduling is used when multiple tasks must be performed and the tasks all have different start and stop times and resource requirements. This makes it more difficult to create a schedule that does not violate any satellite resource constraints, because of the many combinations that the tasks can be ordered in. Heuristic scheduling is a method that attempts to create a feasible schedule by prioritizing the ordering of tasks by employment of one or more heuristics. Some examples of heuristics are scheduling the most valuable task first or scheduling the time slot that is most in demand first. Some heuristic schedulers work with a repair algorithm. These heuristic schedulers work by first very quickly producing a schedule without applying any constraints. The scheduler then works to repair the schedule in all the places where the constraints have been violated. The Hubble Space Telescope employs a heuristic scheduling algorithm called SPIKE that uses this method [25].

Optimal scheduling attempts to create a schedule that is optimal for some defined objective, such as maximizing science observations or minimizing fuel usage. Because of the difficulty in achieving an optimum schedule, few satellites use this method. Some satellite schedules are near optimal, in that the optimizing routine is applied to the schedule for some amount of time, during which the schedule approaches optimal. After a length of time the optimization routine is terminated and the schedule that is last generated is used.

Previous Planning and Scheduling Work Applied to the Integrated Planner

The integrated planner utilizes several concepts from the above planning and scheduling techniques. First, the integrated planner utilizes the Draper EPOS 1.0 optimal planner, which creates an optimal plan and schedule for satellites to observe targets. This plan creates the basic plan that the integrated planner uses. Second, the integrated planner employs a reaction planner to respond to unexpected events that invalidate the plan produced by the optimal planner. The reaction planner re-tasks satellites to regain lost observation time, repairing the plan in terms of observation time regained. This is similar in function to the iterative repair technique that is employed by ASPEN, as a new plan and schedule is produced after the occurrence of an unexpected event.

3.2 Review of Optimal Planning Technique Applied to Earth Observation Problem

Draper Laboratory is currently developing a system that applies optimal planning techniques to the problem of tasking groups of satellites to cooperatively observe various Earth based phenomenon. The system that is currently being developed, called the Earth Phenomena Observation System, or EPOS 1.0, is capable of generating optimal plans for multiple, heterogeneous satellites in a group tasked with the objective of observing multiple targets over a finite planning horizon. EPOS 1.0 commands multiple orbital maneuvers, or burns, to change orbits for the purpose of placing the satellites in orbits that will increase the amount of time that a given target can be observed. The optimal planner utilizes both an orbital mechanics model and a satellite sensor model, which allows a physical representation of realistic satellite movement and observation characteristics to be coupled to the optimal planner.

EPOS 1.0 produces a set of observation plans and maneuver plans for each satellite. Results of simulations are visualized through a combination of data output, charts and the simulation capabilities of a commercially available software package called Satellite Toolkit, or STK.

EPOS 1.0 determines how multiple, heterogeneous satellites can be cooperatively tasked to optimally provide observations of multiple, heterogeneous targets over a defined planning horizon. The optimal planner that is employed by EPOS 1.0 will be described below. Included in this description will be an overview of the concept of operations used by EPOS 1.0 and a mathematical formulation of the problem that is solved by the optimal planner developed for EPOS 1.0.

Overview of EPOS 1.0 Concept of Operations

EPOS 1.0 is designed as a hierarchical mission planner. This hierarchical planner is composed of several levels, as illustrated in Fig. 3.1, that range from the highest levels of system planning through group planning and individual satellite planning down to the lowest levels of planning and control of individual satellite instruments and sub-systems. The reason for using such a hierarchical planning approach is to take a complex problem and decompose it into multiple, less complex problems. It is anticipated that these simpler problems will be largely decoupled and thus can be solved nearly independently [1].

Without decomposition, the problem of determining the optimal group plan of which satellites will view what targets at what time becomes very complex. This is because a group plan that determines when and what size of orbital maneuver a satellite should perform to achieve an observation of the target that helps produces an optimal schedule for viewing the target with all of the satellites becomes intractable as the number of satellites, targets, types of orbital maneuvers possible and mission time frame all increase. If additional parameters are also included, such as communications scheduling between satellites and between satellites and ground facilities, sensor tasking for each satellite and attitude control of each satellite, then the number of parameters that must be planned for and controlled becomes even greater.

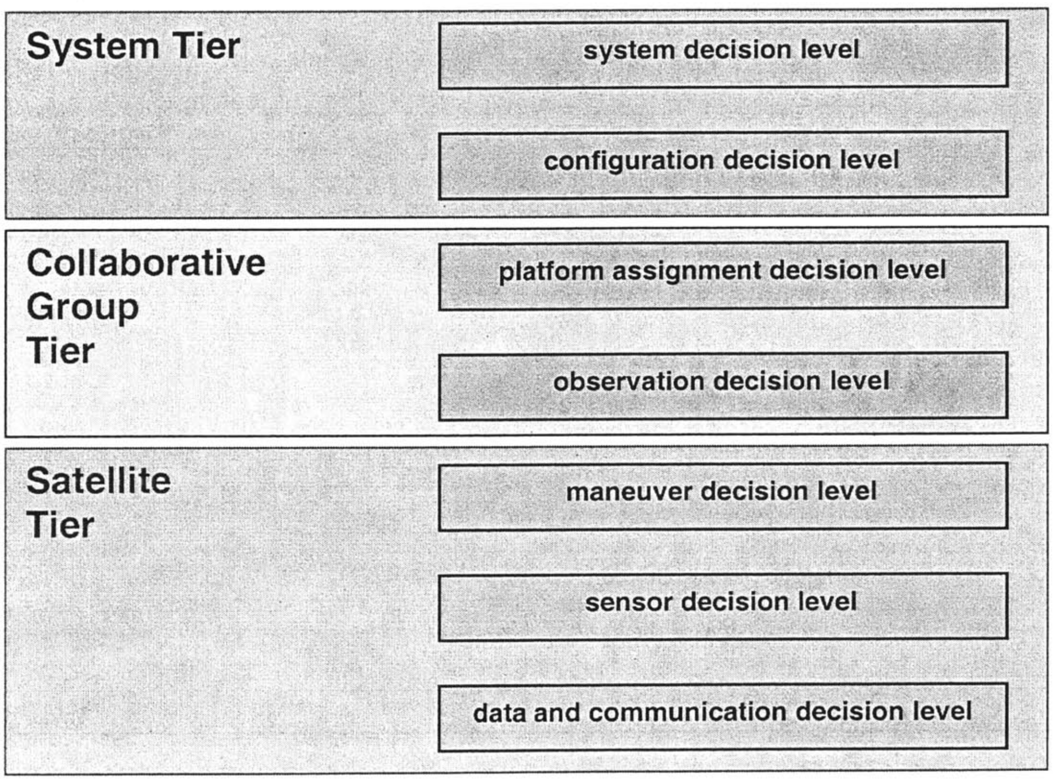


Fig. 3.1. Hierarchical planning architecture used in EPOS 1.0.

At the highest level, System Tier planning, EPOS 1.0 is responsible for efficiently allocating various resources associated with running the entire system, such as centralized control center capabilities and staff, antenna resources and communications bandwidth.

The middle levels pertain to managing the resources available from the satellite group. At the Collaborative Group Tier, data collection tasks are allocated among the various group members and are divided such that the selected targets can be effectively observed. To accomplish these tasks, satellites may be required to perform orbital maneuvers so that they are better positioned in a new orbit that provides an increased amount of observations of the selected target. The Satellite Tier is the lowest level of EPOS 1.0, which provides plans for the various sub-systems onboard each satellite. This planning level allocates the satellite's resources in support of the mission, such as sensor pointing, attitude maneuvers and communications bandwidth. The focus of EPOS 1.0 is on the Observation Decision Level and Maneuver Decision Levels, which are in the Collaborative Group Tier and Satellite Tier, respectively. Planning for these two decision levels can be accomplished either collaboratively at the satellite level or at a centralized location.

Problem Formulation

The problem that is optimized in EPOS 1.0 is to maximize the viewing time that is possible of a given target with a given satellite, while constrained to use no more than a specific amount of fuel. The fuel is used for performing orbital maneuvers to change the satellite's orbital elements. The problem is set up as an acyclical network flow problem. A graphical representation of the network is shown below in Fig. 3.2. The objective is to find the shortest path from the initial state to the end state. The path is a combination of the benefit derived from seeing the target, measured as time in view, and the cost to do so, measured as fuel burned. Each edge has a specific cost and benefit associated with it.

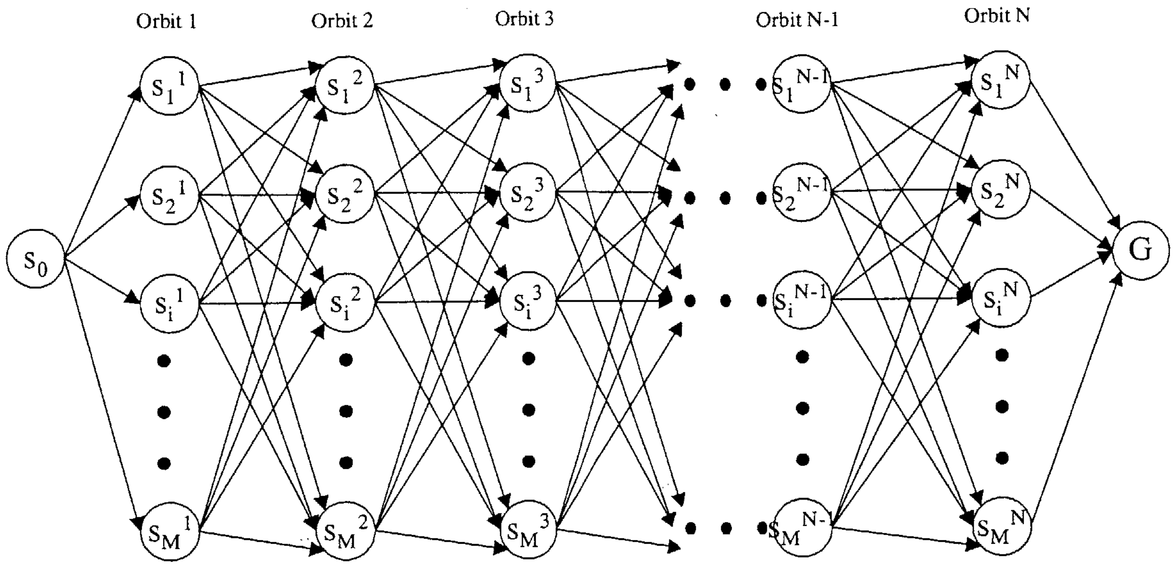


Fig. 3.2. Graphical representation of network flow problem [1].

Each node in the grid, except the initial and end nodes, represents a state that the satellite can be in, which is characterized by a state representation describing its orbit. This state lasts for the duration of the particular orbit that the satellite is in and once the satellite changes orbit it is then represented by a new state. Each state representation is described by the satellite's longitude and satellite orbit period. This is described by the state definition equation below:

$$X = (\theta \quad P) \tag{3.1}$$

The satellite's longitude, θ, is measured once per orbit at a latitude fixed at the equator. The satellite orbit period, P, describes the satellite's repeat ground track and is defined as:

$$P = n / d \tag{3.2}$$

Where equation 3.2 is read as "n days to achieve d orbits". This means that the satellite will revisit the same location after d orbits, which will take n days.

Each edge connecting the nodes represents possible state transitions, that are restricted to occur only once per orbit. Physically, this can represent one of two actions: one, either the satellite is coasting, which means the orbital elements stay constant, or two, there is a phasing burn that the satellite performs that results in a change of its orbital elements. A phasing burn is a type of orbital maneuver and is the only type of burn that is allowed with EPOS 1.0. All nodes have at least one connection coming into and going out of them. This represents the coasting transition, which is always a legal choice for any satellite. This state transition is described as:

$$X_1 \rightarrow X_2 = (\theta \quad P) \rightarrow (\theta - P \quad P) \tag{3.3}$$

Where X_1 is the initial state of the orbit and X_2 is the final state of the orbit. For both states, the orbit period, P, remains constant while the longitudinal position, θ, increments by - P. This is because the satellite is performing a coasting state transition.

Additional connections represent possible changes in the satellites orbit, which is brought about by a phasing burn. All phasing burns used are impulsive, along-track Hohmann transfers. As can be seen from Fig. 3.2, only some nodes are connected. Those that are not connected have a physical meaning that the satellite cannot burn enough fuel to make such a drastic change in orbital elements. This is a constraint placed on the orbital maneuvers that are possible to be performed by each satellite to make the model more realistic, for in practice satellites will only have a limited amount of fuel that they can burn at any one time. When a satellite undergoes a phasing burn, the state transition is described as:

$$X_1 \rightarrow X_2 = (\theta \quad P) \rightarrow (\theta - h(P, P^*) \quad P^*) \tag{3.4}$$

Here, θ and P have the same representations as in coasting transitions, but because the state transition also includes a phasing burn, information on how the phasing burn affects the longitude is required, as represented by P*. The final longitudinal position of the satellite after completing the phasing burn is determined as a function of both the coasting and phasing orbit period.

The grid can be visualized as being composed as a set number of states, arranged in rows and columns. The number of rows in the grid corresponds to the number of possible physical locations that the satellite could be in while over the equator. Physically, this is a continuous set, but has been discretized into steps so that it can be used in a discrete optimization formulation. The orbital position of each satellite type has been discretized into a number of steps of size no greater than 1 deg. The actual size of the discretization depends on the size of the satellite sensor's footprint. Satellites with large footprints only need a relatively course discretization, while satellites with a smaller footprint will need a relatively fine discretized step. For example, with a 1° grid spacing, there would be 360 nodes in each column, for a 0.5 deg grid spacing, there would be 720 nodes in each column.

Each column corresponds to one orbital period of a particular satellite. The number of columns is equivalent to the number of orbits that the user is interested in optimizing over, with burns only being allowed at perigee. Each column represents a particular orbit at a particular time. Once a time period has passed, only orbits in the following time period, located in the adjacent column, are valid, making the network acyclical. So, for example, if the user were interested in a planning horizon for the time equal to 100 orbits, there would be 100 columns. And if there were 360 possible physical locations to choose from, corresponding to the afore mentioned grid spacing of 1°, the total number of nodes, excluding those at the start and end, would be 100 * 360.

The edges connecting the nodes, or states, represent the allowable transitions between states. Each node has at least one edge entering and leaving it, which corresponds to the coasting transition, which is always legal. Additional edges come from allowable phasing burns that change the orbital elements of the satellite's orbit. The transition from one node to the next is governed by allowing the satellite to only perform burns that will result in pre-determined, circular repeat ground track orbits. While burns of any size could theoretically be performed, because of practical limitations imposed by the use of fuel costs of performing burns, burns were constrained to this sub-set. The sub-set of repeat ground track orbits that were specified were chosen so that if no additional burns were made, and the satellite remained in the same orbit, the satellite could still see a target on a regular basis. This means that if a satellite is over the target at least once, multiple viewing opportunities will be available in the future at no additional fuel cost.

The initial state, represented by the single node on the left, is used as a "jumping off point" and the end state, represented by the single node on the right, is used primarily for bookkeeping, and is a dummy end state.

Each edge has a corresponding cost and benefit associated with it. The cost represents the amount of fuel that is used in making the burn in the state transition. If the satellite only coasts from state to state, then the associated fuel cost is zero, as no burn is being performed. The benefit, or viewing benefit, represents the amount of viewing time that that particular satellite will have of a target, during that state, if it makes a specific state transition. If the satellite will not be able to see the target, then the benefit is zero.

In EPOS 1.0 a satellite's observation plan is optimized with respect to one target. Any additional observations that the satellite can achieve for other targets or any observations that coincide with those performed by another satellite do not affect the observation plan created by the optimal planner. In EPOS 1.0 this information is provided to the user for all satellites and can affect the final selection of which satellites will be tasked to view which targets, but this is not done autonomously.

The problem that must be solved is for a given number of states and state transitions, what is the path through the graph that gives an optimal balance of the maximum benefit and the minimum cost possible. This problem can be formulated as a network shortest path (or longest path) problem with side constraints. The formulation that was developed is expressed below in Table 3.1.

Table 3.1. Network shortest path formulation.

$$\max \sum_{i=0}^{M}\sum_{j=0}^{M}\sum_{k=1}^{N} r_j x_{ijk} \tag{3.5}$$

Subject to:

$$\sum_{j=1}^{M} x_{0j1} = 1 \tag{3.6}$$

$$x_{0i1} - \sum_{j=1}^{M} x_{ij2} = 0 \qquad i = 1...M \tag{3.7}$$

$$\sum^{M} x_{jik} - \sum x_{jik+1} = 0 \qquad i = 1..M, k = 2...N-1 \tag{3.8}$$

$$\sum_{j=1}^{M} x_{jiN} - x_{iG} = 0 \qquad i = 1...M \tag{3.9}$$

$$\sum_{i=1}^{M} x_{iG} = 1 \tag{3.10}$$

$$\sum_{i=0}^{M}\sum_{j=1}^{M}\sum_{k=1}^{N} f_{ij} x_{ijk} \leq F \tag{3.11}$$

$$x_{ijk} \in \{0,1\} \tag{3.12}$$

In the above formulation, the objective is to maximize the reward function, as presented in equation 3.5. The amount of viewing time is represented as the reward, r, and x is a binary, integer decision function that represents a legal state transition, with the value of one being a viable transition and zero being otherwise, as presented in equation 3.12. This function is maximized by finding the total reward over N orbits that is possible by transitioning from M possible states to M possible states, minus state transitions that are not viable. The indices represent transitions from state i to state j while in orbit k.

Equations 3.6, 3.7 and 3.9, 3.10 represent the transitions from and into the initial state and goal state, respectively. These equations mean that it is possible for only one transition to occur from the initial state S into the network and from the network into the goal state G. Equation 3.8 ensures that a transition from one to another can be achieved in reverse. That is, if a transition from state 1 to state 2 is possible, that can then be followed by a transition from state 2 back to state 1. Equation 3.12 is a side constraint on the network flow problem that constrains the total fuel amount used throughout the mission to remain at or below a level F. The variable f is a cost function that denotes the amount of fuel required to travel from state i to state j.

Because the network shortest path problem with side constraints is computationally difficult to solve and an eventual goal of EPOS 1.0 is to generate plans in real time, an alternative formulation was developed. This formulation removed the side constraint posed by ensuring that the fuel remained below an absolute limit, F, and shifted it into the objective function in a slightly different form. The new formulation makes use of Lagrangian relaxation, by "relaxing" the fuel constraint and adding it into the objective function. The ensuing formulation that was developed is as follows.

Table 3.2. Network shortest path with Lagrangian Relaxation formulation.

$$\max \sum_{i=0}^{M}\sum_{j=0}^{M}\sum_{k=1}^{N} r_j x_{ijk} - \lambda \sum_{i=0}^{M}\sum_{j=0}^{M}\sum_{k=1}^{N} f_{ij} x_{ijk} \tag{3.13}$$

Subject to:

$$\sum_{j=1}^{M} x_{0j1} = 1 \tag{3.14}$$

$$x_{0i1} - \sum_{j=1}^{M} x_{ij2} = 0 \qquad i = 1 \ldots M \tag{3.15}$$

$$\sum^{M} x_{jik} - \sum x_{jik+1} = 0 \qquad i = 1 \ldots M, k = 2 \ldots N-1 \tag{3.16}$$

$$\sum_{j=1}^{M} x_{jiN} - x_{iG} = 0 \qquad i = 1 \ldots M \tag{3.17}$$

$$\sum_{i=1}^{M} x_{iG} = 1 \tag{3.18}$$

$$\sum_{i=0}^{M} \sum_{j=1}^{M} \sum_{k=1}^{N} f_{ij} x_{ijk} \le F \tag{3.19}$$

$$x_{ijk} \in \{0,1\} \tag{3.20}$$

The objective function is now a network shortest (or longest) path problem of the form:

$$\max[\alpha * benefit - \lambda * fuel] \tag{3.21}$$

where λ is a Lagrangian multiplier that determines the cost of the fuel and α is a scaling factor that allows the difference to be taken between fuel and benefit, meaning that α is both a unit and scale conversion. Small λ's mean fuel is cheap while large λ's make burning fuel prohibitively expensive. This formulation can be solved over a range of λ's to find solutions that generate an acceptable amount of fuel use. This is operationally possible, as there are efficient algorithms available. The constraints in the remainder of the formulation are identical to those discussed above.

The optimal planner that was developed for EPOS 1.0, as described above, is extensively used in this thesis. The integrated planner that was developed combines this optimal planner and a reaction planner. The optimal planner is used in two functions. First, the optimal planner creates an observation plan for each of the satellites. After inserting unexpected events into this plan that cause a loss of observation time, the reaction planner is used to re-task satellites to regain the lost observation time. After a satellite has been re-tasked, the optimal planner is called again. The second use of the optimal planner is to create a new observation plan for the satellite that was re-tasked for the time remaining in the mission.

3.3 Review of Behavioral Planning Technique Applicable to Earth Observation Problem

A behavioral planning algorithm called ALLIANCE was previously developed at MIT's Artificial Intelligence Laboratory [19]. ALLIANCE was designed specifically to address the issue of fault tolerance within a group of cooperative, heterogeneous agents. This was accomplished by making the planning algorithm completely decentralized, which allows each agent to determine what actions need to be completed on their own without the presence of any central planning capability. A learning module was also developed to work with the basic behavioral planning algorithms, called L-ALLIANCE. The learning module allows all agents to update their knowledge about all agent's performance capabilities and to take that new knowledge into account when making future decisions on which agents should be assigned which tasks. This section provides a description of how the ALLIANCE behavioral planning algorithms and L-ALLIANCE algorithms are designed, by presenting an overview of the concept of operations, a brief discussion of the algorithms utilized and an explanation of the mathematical model driving the algorithms.

3.3.1 Description of ALLIANCE

The problem that ALLIANCE is designed to solve is one of how to coordinate a group of agents' planning in a manner that increases the probability that a given set of mission requirements is successfully completed, or adding robustness to the mission. While multiple methods exist that have the aim of raising mission success, the method that the ALLIANCE algorithms focuses on is one of fault tolerance, reliability and adaptability for the group. The problem that ALLIANCE deals with is how to develop a technique that will allow a group to effectively exhibit these characteristics when dealing with total or partial failures of members of the group, operational failures at the group level and the occurrence of unexpected events in the operating environment.

ALLIANCE is a behavioral planning algorithm that is designed to add mission robustness. This is accomplished by making the ALLIANCE behavioral planning algorithms completely decentralized. This decentralization of the planning algorithms allows the group to continue performing a mission even in the event of a failure with one or more members of the group. It is assumed that there is no centralized store of knowledge and that each agent has all the information that is necessary about its own capabilities, the capabilities of other group agents and of the mission. This enables each agent to make decisions about its own actions without the use of one coordinating planner. Information used to make decisions is supplied to each agent at the outset of the mission and throughout the mission. At the outset of the mission each agent receives information concerning characteristics related to the group and the mission. During the mission the agents communicate with one another at predetermined intervals to update the status of each agent in the group and of the status of the mission.

Overview of ALLIANCE Algorithms

ALLIANCE is an architecture that allows for planning and execution of tasks for groups of heterogeneous agents that must cooperatively work together by employing behavioral planning techniques. The ALLIANCE algorithms were specifically designed with the desire for the group of agents to be fault tolerant and reliable, while the group and individual agents are meant to be adaptable. This means that if one or more agents in the group fails to perform its assigned task in the desired manner, the remainder of the group can alter its activities, enabling the group as a whole to still accomplish the mission. This adaptability is at the agent level, which translates into the entire group being adaptable. It is this adaptability that allows the group to respond to faults, increasing the probability that the group as a whole will successfully complete the mission, making the group more reliable. This is achieved by making the ALLIANCE architecture fully distributed over all the agents comprising the group, which allows all agents to be fully autonomous and enables them to continue to perform when other agents experience a failure. ALLIANCE is designed to primarily work with tasks that are loosely coupled and which may have ordering dependencies. This, in certain situations, describes the problem of performing Earth observations with groups of satellites. For some situations viewing different targets may be uncoupled with one another or viewings of the same target may be coupled in time. This means that certain sets of observations from prescribed satellites are desired in a certain order.

Below is a short explanation of each of the major aspects of the ALLIANCE architecture.

Behavioral Planning

The ALLIANCE algorithms make use of behavioral planning techniques. This type of planning technique embeds each agent with a changing "desire" to engage in different activities. This desire factor is called the motivation. The motivation that each agent possesses changes throughout the length of the mission and is a reaction to events transpiring around each individual agent. The ALLIANCE algorithm has two types of behaviors and both affect the level of motivation that the agent possesses. The two types of behaviors present in the ALLIANCE algorithm are impatience and acquiescence. Each one of these behaviors is applicable to each task that a particular agent is capable of performing. Impatience increases an agent's motivation to begin or continue working on a task, while acquiescence decreases or eliminates an agent's motivation to begin or continue working on a task. Each of these behaviors is variable with time and depends on a combination of the agent's actions, of other agents actions and the status of the mission.

A brief, qualitative example is provided to illustrate how these two behaviors affect the motivation of an agent. This example will consist of two agents, A1 and A2, and two tasks, T1 and T2. At the start of the mission both agents know that both tasks must be accomplished to successfully complete the mission. When the mission begins neither agent has started working on any tasks and their initial motivation level is zero. As time progresses, however, and the agents see that no task is being completed, the impatience level of both agents increase. This translates into an increasing motivation to begin one for the required tasks. Eventually, the impatience for one or more of the agents will build to such a point that that agent's motivation will exceed a predetermined threshold and that agent will begin working on a task.

If the impatience of A1 is high for T1, and A2 is not already working on this task, A1's motivation for working on T1 increases. It will continue to increase until the threshold is reached, at which point A1 will begin working on T1. If A2 is idle and observes A1 working on T1, but observes that A1 is taking too long in accomplishing T1, then A2's impatience level will increase until a threshold is reached. If A1 agrees with the assessment that it is not completing T1 in an adequate amount of time, it will "grow frustrated", eventually crossing a threshold on the acquiescence behavior resulting in A1 quitting T1. If A1 is successfully completing a task A2 will begin work on a different task, such as T2.

A mathematical description of the behavioral planning algorithm is described later in this section.

The behavioral planning algorithm is used as a means of recovering after the advent of an unexpected event negatively impacts the original plan. The behavioral planning algorithm is decentralized and does not require that a central control center be notified of the unexpected event or that satellites must wait to react to recover from the event on the directions of a central control center. This is desirable if the central control center takes a long time to generate a new group plan or if the central control center cannot communicate with all satellites. In some applications it may be infeasible to even have a central control center. The behavioral planning algorithm is one method that the satellites can respond to the unexpected event with only a short time delay. It allows all satellites to react to the event based on information about the event that is gathered and distributed by one or more satellites and is then used by all satellites, without each satellite having to identify and monitor the event itself. This means that each satellite need only accomplish the task that is has been assigned, communicate the status of the task to the remainder of the group and listen to the status reports from the rest of the group. If the task can not be completed or the satellite cannot communicate, the rest of the group can infer that the task has not been completed. This allows the group to react without monitoring the status of all the events that other group members are assigned to.

Heterogeneous Groups

A key aspect of a group's utility is that the group will possess all the functions required to complete a mission, but each agent will only possess some of the required functions. This results in the group being composed of multiple, heterogeneous agents. The ALLIANCE algorithm assumes that each agent knows what its capabilities are, is able to determine what capabilities are needed for a task, and can determine if it is able to accomplish the task with its own capabilities. As different abilities are needed at different times throughout the mission, each agent is expected to be able to determine if the particular functions that they provide can accomplish the task. Agents that meet these requirements will allow their motivation to work on a specific task to increase when they are idle and capable of performing that task. If the agent is not capable of performing the task, it takes itself out of consideration of eligible agents and leaves the task to the other agents.

The applicability of ALLIANCE to heterogeneous groups is critical as it is expected that any satellite group will be composed of several types of satellites, each working cooperatively to observe the target. The abilities of each satellite must be know to all other group members so that decisions can be made at the individual satellite level.

Fault Tolerance

The ALLIANCE algorithm is designed to allow the group of agents the ability to autonomously adapt to unexpected faults during the mission, allowing the remaining agents to complete the remaining tasks. Each agent is able to determine if it is not performing well and if other agents are not performing well and to change its behavior accordingly. This is accomplished through a coordination of the impatience and acquiescence behaviors. If an agent feels it is working on a particular task and it begins to take a longer amount of time than it was expecting to finish the task, the agent assumes that it is not performing that task well. In this circumstance, its willingness to quit that task increases. When it increases enough to cross a threshold in the acquiescence behavior, the agent will stop work on that task and notify the other group members that it has ceased work. If an idle agent senses that another agent is taking too long to complete a task or if it is not able to determine the progress that a working agent is making, it assumes that the working agent is not performing adequately. If this happens, the idle agent's impatience increases until it crosses a threshold in the impatience behavior and will then begin the task itself. As the ALLINACE algorithm is fully distributed, each agent can act without intervention from the rest of the group.

The fault tolerance of ALLIANCE is the reason that the algorithms were chosen for the Earth observing satellite problem. If an unexpected event occurs, it will likely impact the ability of the group to complete the mission. The group must be able to re-plan after an unexpected event to regain lost observation time. The fault tolerance of the ALLIANCE algorithms looks at an agent's ability to complete a task. For the Earth observing satellite problem a satellite may not be able to complete a task if either the satellite experiences a failure or the task changes. An example of the second reason may be that cloud coverage prevents a satellite with visual sensors from viewing the target. In this example, the satellite has not experienced a failure with the system, but still cannot accomplish the task. This type of unexpected event does not have a failure that other satellites can monitor directly. The ALLIANCE algorithms treat a failure in the satellite system or a change in the nature of the task similarly.

Fully Distributed Planning

The ALLIANCE algorithm is designed so that all agents have the same source of information on each of the other agent's abilities and performance characteristics and requirements of the mission. This allows each agent to determine what course of action they should engage in, without requiring a central planning source. Agents are required to communicate with the rest of the group, enabling updated information to be shared between agents as the mission progresses. While this leads to much redundant information needing to be made available at the individual agent level and a lack of optimized group behavior, the fully distributed system is very fault tolerant and allows for the group to continue with the mission in the event of one or more agent failures.

The fully distributed nature of the ALLIANCE planning algorithms allows for the group to react in the event of a failure of any one satellite or ground control unit. It is expected that the Earth observation problem will require continuous attention to planning, as new targets appear, targets disappear, targets change or move, satellites are added to the system, satellites experience complete or partial failures or the targets that of are interest change. In the event of a loss of any central control facility, the ALLIANCE algorithms would provide the satellites with a degree of autonomy to make decisions by themselves.

Mathematical Model of the ALLIANCE Algorithm

Below is a description of the mathematical framework that governs how a group of agents will act to accomplish a set of tasks that must be completed in a mission. Recall that agent's actions are governed by two different motivational behaviors, these being impatience and acquiescence. The level that each motivational behavior is at is determined by parameters related to the requirements of the mission, the activities of other agents, the current environment, and the agent's own internal state. The respective levels of both motivational behaviors determine the actions that agents will engage in. For example, if one agent in a group is not accomplishing an assigned task, the impatience that other agents will have for that task will increase, as all the other agents will feel that a mission critical task is not being accomplished. If the impatience level for that task reaches a critical threshold, then a different agent may start working on that task.

Low level behaviours governing primitive survival activities are not included in this model, but are assumed to exist.

The mathematical model used to determine an agent's actions is presented below. The mathematical model is presented in terms of; the set of parameters composing the model, the form of the motivational equation and the actual equations and parameters used to calculate the motivation for each agent.

ALLIANCE Model Parameters

The ALLIANCE mathematical model can be stated as being composed of the following parameters:

- Set of i heterogeneous agents, R_i
- Set of m tasks to be accomplished, T_m
- Set of k actions available to agents to perform tasks, $a_{i1}, a_{i2} \ldots a_{ik}$
- Set of n ways agents can accomplish tasks, $h_i(a_{ik})$: returns task in T that agent r_i is working on when activating action a_{ik}

This information is then coupled with the behaviours of impatience and acquiescence to develop motivation for each agent to perform needed tasks. The motivation for each agent is a time varying function. The motivation for each agent for each task is updated at each time step. The motivation equation for each agent is of the form:

$$m_{ij}(0) = 0 \qquad (3.22)$$

$$m_{ij}(t) = [m_{ij}(t-1) + \text{impatience}] * \Pi(\text{impatience and acquiescence checks})$$

Where,

the current agent is i,

the current task is j,

the current time is t, and

t = 0 is at the start of the mission.

Equation 3.22 initializes the motivation level of each agent for each task to zero at the start of the mission. The motivation level is then recalculated at each time step for each agent and each task, where each agent has a separately calculated motivation level for each task that it is eligible to perform. The motivation level to perform a task is increased by the impatience level. This will mean that an agent with a high level of impatience will reach a motivational state that will allow it to begin a task before an agent with a lower impatience level. If at any time the agent fails an impatience or acquiescence check, for reasons that will be explained later in this section, the motivation level of the agent for that particular task is reset to zero, delaying the time when that agent will start the task.

The actual motivation equation used in the ALLIANCE algorithm is:

$$m_{ij}(t) = [m_{ij}(t-1) + \text{impatience}_{ij}(t)] * [\text{sensoryFeedback}_{ij}(t) * \text{activitySuppression}_{ij}(t) * \text{impatienceReset}_{ij}(t) * \text{acquiescence}_{ij}(t)] \tag{3.23}$$

Where,

- $\text{sensoryFeedback}_{ij}(t)$: determines if activity a_{ij} is applicable for agent r_i at time t
- $\text{activitySuppression}_{ij}(t)$: inhibits other behaviors in r_i when agent r_i is working on a task, meaning that agent r_i can not work on two tasks simultaneously
- $\text{impatienceReset}_{ij}(t)$: impatience of an agent is reset if a different agent begins a task
- $\text{acquiescence}_{ij}(t)$: the agent stops work on a task if one or more other other agents take over or if the agent determines it is not effective in accomplishing the task

Each of the terms above in 3.22 that determine the level of motivation for an agent must be calculated at each time step. Several additional terms and parameters that are used in this calculation are presented and explained below, starting with the agent's level of activity activation.

Activity Activation

Agents begin an activity when their motivation level for that activity reaches or exceeds a predetermined threshold of activation. This threshold of activation is defined by the parameter, θ, and is constant for all agents and tasks throughout the mission. While it may seem that θ should be a variable value, dependent on what agent and what task are being examined, a constant θ is acceptable because other parameters related to impatience and acquiescence, which will be discussed later in this section, adequately represent the variability of different agent's status and capabilities with respect to the mission tasks given.

Inter-Agent Communications

An important parameter that does not appear directly in the motivational equation expressed in 3.22 but is required for determining several other parameters is the inter-agent communication parameter, comm_received. Inter-agent communications are assumed to take place at some predetermined frequency and serve the purpose of distributing information concerning the status of task completion and group activities to each agent. The expression for comm_received is given below:

$$comm\_received(i,k,j,t_1,t_2) = \begin{cases} 1 & if \quad r_i \; received\; message\; from\; r_k\; concerning \; task\; h_i(a_{ij})\; in\; time\; span\; (t_1,t_2) \\ 0 & otherwise \end{cases}$$

Two additional parameters define an agent's communication characteristics, these being ρ_i and τ_i, which, respectively, denote the rate at which agent i will broadcast its activities to the group and the time that agent i will wait without receiving a communications update from an agent before determining that agent is no longer functioning.

Impatience

Impatience is the driver behind increases in an agent's motivation level. The rate of impatience varies across agents, tasks and time. The value of an agent's impatience level with a tasks at a given time should represent that agent's understanding of how appropriate it is for performing a certain task, given the importance of that task to the overall missions and the current activities of all agents in the group. If the agent is currently unoccupied and believes that it could better accomplish a certain task, its impatience level will rise and will cause a corresponding significant increase in that agent's motivation level to begin working on a task. If the agent believes that it is not the best agent suited for a particular task or believes that an agent currently working on the task is doing an adequate job, then that agent's impatience level will be small and will correspondingly cause only a small increase in the agent's motivation to begin working on that task. If the agent is currently involved in a task, it's impatience level for other tasks will be zero and will therefore not cause a rise in the agent's motivation to begin a different task.

Three parameters are defined to determine an agent's level of impatience. These parameters are $\delta\_fast_{ij}(t)$, $\delta\_slow_{ij}(k, t)$ and $\phi_{ij}(k, t)$. The first two parameters define the rates of impatience that each agent will have for each task over the course of the mission. As defined, $\delta\_fast_{ij}(t)$ will cause an agent's motivation to rise at a faster rate than $\delta\_slow_{ij}(k, t)$. The third parameter defines the length of time that an agent will allow its actions to be determined by the communications sent from other agents in the group. For example, an agent will have a low impatience level if it continues to receive communications from other agents stating that they are making satisfactory progress on a task for only a certain amount of time. After that amount of time, if the agent believes that a task should have already been completed its impatience level will increase, even if it is continuing to receive messages from other agents stating that they are working on the task in a satisfactory manner. More precisely, the three parameters can be interpreted as:

- $\delta\_fast_{ij}(t)$ – the impatience level that agent i will have for task j at time t,
- $\delta\_slow_{ij}(k, t)$ – the impatience level that agent i will have for task j if an agent k is working on task j at time t, and
- $\phi_{ij}(k, t)$ – the time that agent i will allow agent k's communication broadcasts to influence its impatience for a task j at time t.

These parameters combine to determine the impatience level in the following manner:

$$impatience_{ij}(t) = \begin{cases} \min_k\left(\delta\_slow_{ij}(k,t)\right) & \text{if} \quad comm\_received(i,k,j,t-\tau_i,t) = 1 \quad and \\ & \quad comm\_received(i,k,j,0,t-\phi_{ij}(k,t)) = 0 \\ \delta\_fast_{ij}(t) & otherwise \end{cases}$$

The above equation sets agent i's impatience level to a low level if it has received broadcasts from other agents stating that they are accomplishing task j. As long as the communication broadcasts are received at a frequency greater than τ_i and have not exceeded a time $\phi_{ij}(k, t)$, then the impatience level will be set low enough to give agent k adequate opportunity to accomplish task j. If either of these two constraints are violated, meaning that either agent i has not received communications from agent k in a time period greater than τ_i or that agent k has been working on task j for a time period longer than $\phi_{ij}(k, t)$, then agent i's impatience level will be set to a higher rate.

One other parameter influences an agent's impatience level, this being the impatienceReset parameter. If an agent receives a communication from a different agent that it has started a task for the first time, its impatience will be reset to zero. This allows a different agent to either start or take over a task and have sufficient time to try and complete the task before another agent tries to take over. The impatienceReset parameter can be described as:

$$impatienceReset_{ij}(t) = \begin{cases} 0 & \text{if} \quad \begin{array}{l} comm\_received(i,k,j,t-\delta t,t) = 1 \; and \\ comm\_received(i,k,j,0,t-\delta t) = 0 \\ where\ \delta t\ is\ the\ time\ since\ last\ communications \end{array} \\ 1 & otherwise \end{cases}$$

The above equation sets the motivation for agent i to zero if a communication check stating that agent k has started task j in the last δt time steps has occurred.

Acquiescence

The parameter of acquiescence partially controls when an agent quits working on a task and relinquishes the task to a different agent. This is different than if an agent successfully completes a task and then stops working on it. An agent will acquiesce a task under different conditions when either there is another agent to take over the task or if the agent believes that it is not going to be able to successfully complete the task. These two characteristics are defined by the parameters $\psi_{ij}(t)$ and $\lambda_{ij}(t)$, respectively. More precisely, these parameters can be interpreted as:

- $\psi_{ij}(t)$ – the length of time that agent i will maintain activity on task j before yielding to another agent, and
- $\lambda_{ij}(t)$ – the length of time that agent i will maintain activity on task j before giving up and moving on to a different task.

These parameters combine to determine an agent's willingness to acquiesce in the following manner:

$$acquiescence_{ij}(t) = \begin{cases} 0 & \text{if} \quad \begin{array}{l} r_i \; has\,been\,working\;on\,task\;j \geq \psi_{ij}(t) \quad and \\ comm\_received(i,k,j,t-\tau_i,t) = 1 \\ or \\ r_i \; has\,been\,working\;on\,task\;j \geq \lambda_{ij}(t) \end{array} \\ 1 & \qquad\qquad otherwise \end{cases}$$

The above equation sets the parameter acquiescence equal to zero, which in turn sets the motivation for agent i equal to zero if the agent has either, one, been working on a task for longer than a period $\psi_{ij}(t)$ and has received a communication from agent k in the last τ_i time steps or two, has been working on the task for longer than $\lambda_{ij}(t)$ time steps.

Sensory Feedback

The sensory feedback parameter is used to represent whether a certain set of actions are acceptable to complete a task. This means that if a task needs to be completed, each agent must determine if it has the appropriate tools to complete the task at a given time. If the agent determines that it can not perform the needed action to complete a task, the parameter sensoryFeedback is set to zero, which in turn sets the motivation level for the agent to zero for that particular task. This is expressed as follows:

$$sensoryFeedback_{ij}(t) = \begin{cases} 1 & \text{if } actions\ available\ to\ r_i\ are\ applicable\ to\ task\ j\ at\ time\ t \\ 0 & otherwise \end{cases}$$

Activity Suppression

It is assumed that each agent can only accomplish one task at a time. With this assumption, after an agent begins a task, its motivation level and desire to start any other task drops to zero for the remainder of the time that it is working on a task. This is captured in the parameter activitySuppressionij(t) and expressed as:

$$activitySuppression_{ij}(t) = \begin{cases} 0 & \text{if } r_i\ is\ working\ on\ task\ x \neq j\ at\ time\ t \\ 1 & otherwise \end{cases}$$

Summary of Parameters

A summary of the parameters discussed above are presented below, along with a brief discussion on how values are chosen for the parameters.

Table 3.3. Summary of parameters used in ALLIANCE algorithms

- ϕ - time that agent i is influenced by other agents' communications
- δ_fast - high growth rate of impatience

- δ_slow - low growth rate of impatience
- ψ - time agent will maintain task before yielding to another agent
- λ - time agent will maintain task before giving up
- θ - level before motivational behavior is activated
- ρ - rate at which agent i broadcasts its activities to other agents
- t - time that agent will allow between communications before assuming teammate has ceased to function

The first eight variables in Table 3.3 are set by the user, often running several test cases in order to determine what the proper variable settings should be for a particular set of tasks. Variable values are set to emphasize performance strengths of each agent. It is important to set these properly as the variables determine for each agent what tasks are accomplished, in what order and with what efficiency. These variable values influence idle time and task re-allocation between agents. Poor choices in variable settings will waste group resources. This is an especially bad problem during missions where the tasks to be accomplished change unexpectedly or the abilities of the agents change unexpectedly (for example a degradation of performance during the course of the mission). The learning module in L-ALLIANCE is designed to take these possibilities into account by continuously updating the variable values with new information on agent performance and mission requirements.

3.3.2 Description of L-ALLIANCE

L-ALLIANCE builds directly upon the framework outlined above for the ALLIANCE algorithms by adding a learning module, which updates the variable values continuously to take into account new information regarding tasks and agent capabilities. L-ALLIANCE's learning capability is derived from continuously comparing the agent's current performance of accomplishing an activity with its own past performance for completing that activity and all other viable activities, as well as comparing its current performance to all other agents' performance in accomplishing that activity. Agent performance is measured in time required to complete the task. This comparison of the agents' performance allows it to influence what activity a particular agent will begin next, choosing the activity based on what is needed for the mission, what activity it accomplishes best in relation to the other agents and what activity it accomplishes best for the range of activities that the agent is able to perform. This comparison results in an agent selecting the activity that it will be best at. This is accomplished by modifying the impatience parameter. When an agent completes a task in a shorter amount of time that was planned for, the agent interprets that as an improvement in performance over what was anticipated. The next time this task is available, the agent will be more impatient to work on the task because of the better than expected performance. Additionally, if the agent is able to perform multiple tasks better than any other agent, from this set of tasks the agent will select the task that takes the longest time to complete. This allows the agent best suited for the activity to be assigned to work on it for the longest period of time.

The purpose for the addition of the learning module to the ALLIANCE algorithms is that it increases the efficiency in which the agents can accomplish the mission. When just using the ALLIANCE algorithms, all relevant parameters that describe each agent's performance capabilities and hence influence which agents get assigned to which task, are determined before the mission begins. If the mission either changes over time once it begins or is different than anticipated at the outset, these parameters will be incorrectly set. The result will be a set of agents that do not accomplish the mission as well as they could or may not be able to accomplish the mission at all. L-ALLIANCE allows the parameters that are set before the mission to be updated automatically throughout the mission with current information. This results in the parameters better reflecting the current state of the mission and the performance capabilities of each agent.

Overview of L-ALLIANCE Algorithms

Below is a short description of the major aspects of L-ALLIANCE. As L-ALLIANCE and ALLIANCE share the same basic algorithms, only differences found between the two algorithms and new capabilities in L-ALLIANCE will be covered below.

Learning Module

The learning module in L-ALLIANCE influences each agent's motivation level through dynamically changing the impatience and acquiescence levels as each agent performs various activities necessary to the mission. Past performance of each agent's performance for each task is taken into account when determining what task an agent will begin and how long it will work at the task before another agent takes over.

Task Categories

All tasks are continuously re-divided into two categories, Category I tasks which are the tasks that an agent does better than any other agent, and Category II tasks, which are tasks that other agents exhibit superior performance. When determining what task an agent will choose, it will try and begin work on a Category I task, if no other agent is working on it and if it is required for the mission to proceed. From the tasks in Category I, the agent will pick the task that it takes the longest to complete. This is to allow the task that takes longest to complete to be begun by the agent that is best able to complete the task, an attempt at a "temporal optimum" group assignment (meaning that it is the optimum assignment for the tasks that need to be completed and the agents available to complete them at that instant in time). This assumes that all tasks are of equal importance.

Boredom

An additional means of assigning agents to tasks is achieved with a "level of boredom" indicator. If an agent does not have any tasks currently available to it, it will become bored. Once a certain level of boredom has been reached, the agent will begin a task even if it knows that another agent can better accomplish the task. This sub-optimal assignment of agent resources cuts down on idle time for the group. An example of this would be if one agent is able to perform most tasks better then the rest of the group, instead of waiting for the one agent to perform all the tasks that it is the best at, the other agents will become bored just waiting and will begin tasks that they are not the best at rather than sit idle.

Strategies

There are multiple strategies for determining how the agents will interact with each other as they learn. What strategy is chosen will determine how each agent's impatience and acquiescence levels will be set, which will greatly influence what tasks get chosen, how long a agent will work on a task before quitting and how often the agents will re-task themselves. Two examples of strategies that could be used are, one, basing all decisions on the performance level of the best agent, and two, basing all decisions on the average performance level of the agents accomplishing the task. Each strategy has benefits and shortcomings. Additional strategies could also be developed.

A brief overview of these two types of strategies is as follows. Strategy 1, basing decisions on the best performing agent, will ensure that agents that do not perform a task as well as the best performing agent will not slow down the mission if they can not complete the task in the same amount of time that the best performer can. If one of these agents begins a task and cannot finish in this time, the impatience level of the other agents will quickly rise and they will shortly take over the task from the current agent. The downside of this is that it is often impossible for the best performing agent to be working on the task of interest, as it may be engaged in a different task. If this is the case, then unnecessary re-tasking of the agents occurs with this strategy.

Strategy 2, basing decisions on the average performance of the agent working on the desired task, will allow an agent that has begun a task to continue to perform the task as long as it is not taking longer than that agent usually takes to complete that task, on average. If the agent begins to take a longer amount of time than it does on average, other agents will begin to have their impatience level rise and will soon re-task themselves to completing that task. The downside to this strategy is that if a better performing agent is available, it will not try and re-task itself to quickly complete the task, preferring to give the agent currently working on it a chance to complete it itself.

Mathematical Model of the L-ALLIANCE Algorithm

Below is a description of the mathematical model that governs how L-ALLIANCE operates. It is based on the ALLIANCE algorithms and shares many of the same variables. Some of the variables are unchanged, some variables that are statically set under ALLIANCE are continuously updated under L-ALLIANCE and some additional variable are required. The prime emphasis of all new variables and modification of variables previously found in ALLIANCE is to incorporate a comparison of past performance to current performance for all agents for all tasks.

The actual motivational equation used in L-ALLIANCE is very similar to the one used in ALLIANCE as presented in Equation 3.22. The only difference between the ALLIANCE and L-ALLIANCE motivation equation is the addition of one additional multiplier, learnedAgentInfluence, and the recalculation of several of the existing terms to include information generated from the learning module. The motivation equation is presented below in Equation 3.23.

$$\begin{aligned} m_{ij}(t) = {} & [m_{ij}(t-1) + \mathrm{impatience}_{ij}(t)] * \\ & [\mathrm{sensoryFeedback}_{ij}(t) * \mathrm{activitySuppression}_{ij}(t) * \mathrm{impatienceReset}_{ij}(t) \\ & * \mathrm{acquiescence}_{ij}(t) * \mathrm{learnedAgentInfluence}_{ij}(t)] \end{aligned} \tag{3.24}$$

Where,

- $\mathrm{learnedAgentInfluence}_{ij}(t)$: determines the boredom level of agent i for task j, if no other agents are working on task j and agent i does not have any Category I tasks to complete
- all other tasks have the same function as described in Equation 3.23

Additional variables that are used to calculate the motivation are shown below. All variables that are needed are listed below. Variables that are calculated in an identical manner as ALLIANCE are not described, but are listed for completeness.

Activity Activation

The threshold for an agent to begin working on a task is identical to that described above in the ALLIANCE section.

Inter-Agent Communication

All parameters related to inter-agent communications are identical to that described above in the ALLIANCE section.

Impatience

The impatience parameter has the same function as that described above in the ALLIANCE section. It is calculated in the same manner as that described above, but the parameters that it is a function of are calculated in a different manner using the L-ALLIANCE learning module. In ALLIANCE, there are three parameters that determine an agent's level of impatience, all of which are set at the start of the mission and remain constant throughout the mission. In L-ALLIANCE all three of the parameters are updated throughout the mission to better reflect increased knowledge of the mission requirements or increased knowledge of an agent's performance capability. The three parameters that determine impatience are presented below.

- $\phi_{ij}(k, t)$ – the time that agent i will allow agent k's communication broadcasts to influence its impatience for a task j at time t. In L-ALLIANCE this time is defined as

  $$\phi_{ij}(k, t) = task\_time_i(x,j,t)$$

  This allows the time to be updated depending on what task and what agent is in communications. The variable x represents an agent and is determined based on what strategy is selected. For example, if Strategy 1 (best performing agent) is used, then x is the agent that has the lowest value of task_time.

- $\delta\_slow_{ij}(k, t)$ – the impatience level that agent i will have for task j if an agent k is working on task j at time t. In L-ALLIANCE this is defined as

$$\delta\_slow_{ij}(k, t) = \theta / \phi_{ij}(k, t)$$

This allows the impatience level to be a function of how long the task is expected to take to complete. Tasks that take a longer time to complete will set the impatience level at a lower rate.

- $\delta\_fast_{ij}(t)$ – the impatience level that agent i will have for task j at time t. In L-ALLIANCE this impatience level is determined in a relatively complex manner that takes into account the task category, current task time, shortest and longest tasks available and permissible delays. In L-ALLIANCE this is defined as

$$\delta\_fast_{ij}(t) = \begin{cases} \dfrac{\theta}{min\_delay + (task\_time_i(i, j, t) - low) * scale\_factor} & \text{for task category II} \\ \dfrac{\theta}{max\_delay - (task\_time_i(i, j, t) - low) * scale\_factor} & \text{for task category I} \end{cases}$$

Where,

o $min\_delay = \text{minimum allowed delay}$

o $max\_delay = \text{maximum allowed delay}$

o $high = \max_{k,j} task\_time_i(k, j, t)$

o $low = \min_{k,j} task\_time_i(k, j, t)$

o $scale\_factor = \dfrac{max\_delay - min\_delay}{high - low}$

From this definition, the impatience level is set to a high state when the task falls into category one and takes a long time to complete. As the time required to complete a category one task decreases, the impatience level is set to a lower state. For category two tasks, shorter tasks are assigned a higher impatience rate. As the length of time required to complete a category two task increases, the impatience level decreases. This is done to show the preference of an agent to work on a task that it is good at accomplishing and takes a long length of time or that it may not be the best at and takes a short time to complete. Agents try to avoid working on tasks that they cannot accomplish efficiently and take a long time to complete.

The impatienceReset parameter is the same as defined above in the ALLIANCE section.

Acquiescence

The acquiescence calculation in L-ALLIANCE is similar to the impatience calculation in that it is accomplished in the same manner as with ALLIANCE, but the parameters that it is a function of are calculated in a different manner. Acquiescence is a function of two parameters, λ and ψ, which control the time that an agent wants to maintain work on a task before giving up to try another task and the time that an agent wants to maintain work on a task before yielding to another agent, respectively. Of the two parameters, only ψ is updated in L-ALLIANCE and depends on the time required to complete a task. In L-ALLIANCE this is defined as

$$\psi_{ij}(t) = \min_{x} \text{task\_time}(x, j, t)$$

Where,

x is a variable representing an agent, depending on the strategy selected. For example, if strategy one is selected, then ψ will acquiesce to another agent after a time equal to the minimum time required to complete the task by any agent has elapsed.

Sensory Feedback

The sensory feedback parameter is identical to that described above in the ALLIANCE section.

Activity Suppression

The activity suppression parameter is identical to that described above in the ALLIANCE section.

Learned Influence

Each agent experience two types of learned behavior when using L-ALLIANCE. These relate to the categorization of tasks based on each agent's performance and the introduction of a boredom parameter. Each of these parameters is explained below.

Tasks are continuously categorized by each agent into either one of two categories. Category one tasks are tasks that an agent is able to complete in a shorter amount of time than any other agent and category two tasks are all other tasks that the agent performs at a level lower than at least one other agent. This is defined as

$$\text{task\_category}_{ij}(t) = \begin{cases} 1 & \text{if } (\text{task\_time}(i,j,t) = \min\limits_{k \in \text{agents present}} \text{task\_time}(k,j,t)) \\ & \textit{and } (\sum\limits_{x \in \text{agents present}} \text{comm\_received}(i,x,j,t-\tau_i,t) = 0) \\ 2 & \text{otherwise} \end{cases}$$

This basically states that a task is placed into category one for an agent if that agent has the lowest time required to complete the task of any agents and no other agents are currently working on the task. If both of these conditions are not met, then the task is placed into category two. Tasks are re-categorized every time an agent is about to begin a new task.

Boredom is expressed in the learned_influence parameter. Boredom is essentially an agents desire to start a task if it is idle, even when it is not good at completing any of the tasks that are currently available. This assumes that it is better for an agent to be working on a task even when it is not effective at accomplishing it. Boredom is defined through the following series of equations.

$\text{boredom\_threshold}_i$ = level of boredom at which agent i begins a task not being completed

boredom_rate_i = rate of boredom of agent i

$$\text{boredom}_i(t) = \begin{cases} 0 & \text{when } t = 0 \\ (\prod_j \text{activity\_suppression}_{ij}(t)) * & \text{otherwise} \\ (\text{boredom}_i(t-1) + \text{boredom\_rate}_i) & \end{cases}$$

$$\text{learned\_influence}_{ij}(t) = \begin{cases} 0 & \text{if } (\text{boredom}_i(t) < \text{boredom\_threshold}_i) \\ & \text{and } (\text{task\_category}_{ij}(t) = 2) \\ 1 & \text{otherwise} \end{cases}$$

Chapter 4

A Behavioral Planning and Learning Algorithm

A behavioral planning algorithm was chosen to be integrated with EPOS 1.0 to provide robustness to the mission to counteract the occurrence of an unexpected event. EPOS 1.0 is a system capable of generating an optimal plan for a group of satellites performing observations on specific Earth based targets, but has no mechanisms for handling unexpected events. This chapter provides an overview of work that was completed in studying ALLIANCE, the pre-existing behavioral planning system, and modifying the ALLIANCE algorithms to increase the applicability to the Earth observing problem. This was accomplished by first developing a simple problem containing functional similarities to the Earth observing satellite problem and then applying ALLIANCE to it. Also included in the chapter is the application of a pre-existing learning algorithm, L-ALLIANCE, to the same functionally similar problem. The chapter is concluded with a discussion of how the behavioral planning and learning algorithms were modified to make them more appropriate for the developed problem, results obtained when applying them in a simulation environment and lessons learned that are applicable when applying the system to the actual Earth observation problem.

This chapter contains an example of how the ALLIANCE algorithm can be applied to a problem in a domain that is different than the Earth observing satellite problem. The purpose of this chapter is to illustrate how ALLIANCE works and how the various parameters interact to yield different group behaviors. ALLIANCE was studied in the waste container movement problem domain because of the absence of the difficult to predict and work with orbital mechanics models that are necessary in the Earth observing satellite problem domain. This allowed ALLIANCE to be better understood before applying it to the Earth observing satellite problem. The work in this chapter does not explicitly allow ALLIANCE parameter values to be set for the Earth observing satellite problem, though general insights as to how the parameters interact aid in setting the parameter values for the Earth observing satellite problem.

The use of a planner that reacts autonomously, by reconfiguring the task distribution for a group of agents, is one means of achieving increased robustness during a mission. The reaction planner is used in the event that a complete or partial failure with one or more agents occurs, or if the environment that the agents operate in changes. One potential application for this type of reaction planner would be to integrate it with a planner that produces an optimal plan for the group and to only engage the reaction planner in the advent of an unexpected event. Such a use would necessitate a thorough understanding of the reaction planner before integrating it with any other system. This can be achieved by applying the chosen reaction planner to a problem that has a lower complexity and possesses functional similarities with the problem of interest. ALLIANCE was utilized to accomplish this function.

The ALLIANCE algorithm was applied to a toxic waste relocation problem. In this problem, a number of agents were created that were tasked with moving a container filled with waste to a secure facility. While the movement of toxic waste and the tasking of satellites to observe Earth based targets appears to have few similarities, there are several important comparisons that can be drawn between the missions generated by the two problems. These comparisons make the study of ALLIANCE algorithms applied to the toxic waste problem of interest. The remainder of this section will cover how the waste movement mission was developed, explain critical similarities and differences between it and the Earth observation problem, provide an overview of results gained from analyzing the simulation created and discuss lessons that are appropriate when applying the L-ALLIANCE learning module and when applied to the Earth observing satellite problem.

4.1 Problem Definition

The following section presents an overview of the mission that was developed to study the ALLIANCE algorithms in greater detail. This includes an expanded explanation of the mission profile, discussions on the agent characteristics that were developed for the waste movement problem, an overview of the environment that the agents must operate in and an elaboration on the types of failures and their corresponding effects on the group during the course of the mission. Rationale on the characteristics of the mission and how the mission relates to the Earth observing satellite problem are discussed in more detail in the following section.

Mission Profile

The mission, created for the waste movement problem, was to coordinate a group of four heterogeneous agents to work together cooperatively, in order to move a large waste container across a room and into a secure holding area. The movement of the waste container and the placement of the container into the holding area required that the agents work together. The goal was the movement of the waste container. The metric that was used to determine the success of the mission after completion was the amount of time that was required. A mission taking longer than some predetermined amount of time that ended with the waste container not being delivered into the holding area was not judged as successfully completed. Missions being completed in shorter amounts of time were seen as using the resources of the group more efficiently than missions that took a longer amount of time to complete, accounting for time delays due to failures and the amount of knowledge of the mission available to the group a priori.

In order to complete the mission successfully, three tasks had to be accomplished. These tasks were opening the door to the holding area to allow the waste container to enter and pushing the box into the holding area. The box was large enough relative to the agents such that it had to be pushed from each end, with pushing on the right and left ends each constituting a separate task. These three tasks were chosen for the following reasons; first, they represented multiple tasks that had to be accomplished to successfully complete the mission. Second, the required tasks were a mix of homogenous tasks (pushing the symmetric box ends) and heterogeneous tasks (opening the door and pushing the box ends). Third, some of the tasks had to be completed multiple times (repeatedly pushing the box ends across the room) while the other task only had to be completed once (opening the door to the holding area). Fourth, there were different levels of ordering dependencies in the mission. While the door had to be opened before the box could be pushed into the holding area, the mission was sufficiently long enough so that there was plenty of time to open the door at any time during the mission, making the time constraint of the ordering dependency very loose between the tasks of opening the door and pushing the box into the holding area. On the other hand, because of the size of the box, it could only be pushed from one end twice before it had to be pushed from the other end, so that its position relative to the direction of travel stayed roughly constant. This ordering dependency still allowed a decision to be made as to which end of the box would be pushed next, but constrained the decision rather tightly over the course of the mission. These tasks were designed in this manner so as to develop a mission that had several functional similarities with the tasks in the Earth observing satellite problem.

Agent and Group Characteristics

A group of four agents with heterogeneous capabilities was created for the waste movement problem. Of the four agents, two were identical, homogenous agents and the other two were of two different types, bringing the total types of agents in the group to three. The agents were differentiated by their ability to move the box and by the speed in which they move in general. Because the mission success was judged based on how quickly the mission could be completed, or minimizing mission time, the faster agents are seen as more desirable. At the beginning of the mission, all agents are assumed to be fully operational and are able to complete any of the necessary tasks, meaning that all agents are capable of pushing either end of the box or opening the door to the holding area. In the course of the analysis, several different scenarios were examined, which included, one, each agent possessing imperfect knowledge of both their own performance capabilities and the performance capabilities for other members in the group and, two, perfect knowledge of the performance capability for all group members, including themselves.

Operating Environment

The operating environment that the agents worked in was affected by a variable terrain profile that affected agent movement and the speed in which the box could be pushed. The terrain profile included a plateau at the start and end of the course and a hill with a variable slope in the middle of the course. The terrain profile is shown below in Figure 4.1. As is intuitive, it took a longer amount of time for the agents to either travel uphill or to push the box uphill. Also, any movement downhill took the same extended time as traveling uphill, as it was assumed that the agents would have a more difficult time navigating downhill, as opposed to speeding up after receiving a "gravity assist". The distance that the box was pushed is the same uphill and downhill.

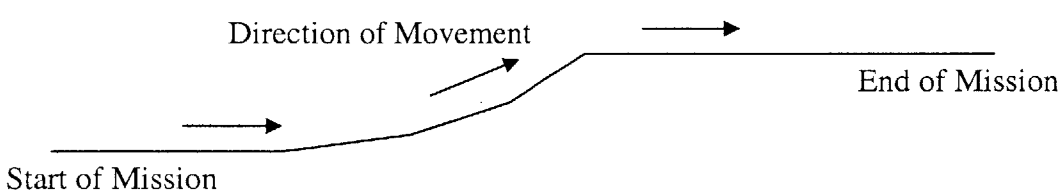


Fig. 4.1. Terrain profile of waste movement mission.

Failures

Two major classes of failures were examined in the waste movement problem, these being complete failure of one or more agents or partial failure of one or more agents occurring at some time over the course of the mission. Complete failure is characterized by the inability of an agent to complete, start or continue any task for the remainder of the mission. Partial failures are characterized by either, one, the inability of an agent to complete, start or continue selected tasks while still being able to perform other tasks with no degradation to its ability or, two, agents possess the ability to still perform all tasks but at a lower level of performance than was possible at the start of the mission. It is assumed that any failure is permanent and that the agents do not experience a restoration of their performance during the mission.

The only means that the agents have for determining what other members in the group are doing is through transmitting and receiving communication updates from each agent. Agents are assumed to be able to monitor their own progress when working on a task and then relaying that progress to the rest of the group through periodic communication updates. It is assumed that the agents do not have the ability to independently assess the status of other agents through passive observation of those agents. For additional fault tolerance, if an agent ever stops receiving communications from a different member in the group over a predefined time limit, it is assumed that the agent that has quit communicating has experienced a failure of some sort. If this occurs, then the remainder of the agents will assume that the non-communicating agent is incapable of performing any tasks and will not include that agent in any future task allocation planning. This includes cases where it either stops communicating entirely or communicates that it is not able to successfully complete a task.

4.2 Functional Comparison between Waste Movement and Earth Observing Satellite Problem

As stated earlier, the waste movement problem was developed to explore and modify the ALLIANCE algorithms in a problem that was less complex but possessed functional similarities to the Earth observing satellite problem. The following section lists and discusses the critical similarities and differences between the waste movement and Earth observing satellite problems. A discussion of the lessons learned in the waste movement problem that are applicable to the Earth observing satellite problem is provided in a following section.

Critical Similarities

The following is a list of how and an explanation of why the two problems are functionally similar.

Multiple agents

A group of agents was deemed necessary for reasons of effectiveness in accomplishing both the problems of waste movement and Earth observing satellites. As was previously explained, the Earth observing satellite problem cannot be effectively completed without a group of satellites. Because the Earth observing satellite problem required a group of satellites, the waste movement problem was created so that a group of agents would be required to work cooperatively together also. Having a group of agents, as opposed to one monolithic agent, means that the groups of agents in both problems must perform some identical additional tasks. Some examples of this are decomposing the problem to be worked on efficiently by multiple agents and effectively coordinating the agents. For coordinating the groups of agents, the same basis reaction planning algorithms were used in both problems. In the Earth observing satellite problem an additional level of planning was performed, but in the advent of a failure, the same reaction algorithms were applied as to the waste management problem.

Heterogeneous and homogenous agents in groups

The agents that comprise the groups in both problems are a combination of multiple homogenous and heterogeneous agents. This means that some of the agents will have the same performance capabilities as other agents and some will have different performance capabilities. This adds a dimension of complexity to both problems as it must be decided which agent is best suited toward accomplishing a given task when planning. In the case of the homogenous agents, multiple agents may have the exact or similar performance capability in performing a task, making the decision of one between two equals, or nearly equals. In the case of heterogeneous agents a decision must be made between differing performance capabilities, which necessitates a clear understanding of the task to be solved and the best application of the abilities that each agent possesses. In the Earth observing satellite problem, targets may need one or more satellite to be properly observed. Sometimes the same type of observations are required, but at an interval that requires multiple homogenous satellites, and in some instances different types of observations are needed, dictating the use of multiple heterogeneous satellites. This need was modeled in the waste movement problem by creating a group of agents that had both similar and differing performance characteristics from one another. While all agents could nominally accomplish all tasks, scenarios were also developed that caused some agents to become partially disabled, which resulted in some agents being able to complete only selected tasks. These partial failures resulted in even greater agent heterogeneity in the waste movement problem.

Multiple tasks

In order to successfully accomplish the waste movement mission, multiple tasks must be accomplished. Having a set of tasks that need to be accomplished, rather than just one task, means that the group of agents must determine not only which agents will be assigned tasks, but what tasks must be completed and in what order. This is similar to the Earth observing satellite problem that has multiple targets that must be observed so that the mission can be completed successfully.

Heterogeneous and homogenous tasks

The set of tasks that comprise the waste management mission are both heterogeneous and homogenous in nature. Homogenous tasks require the same type of activity to complete them, while heterogeneous tasks require a different set of activities. Heterogeneous tasks may require different types of agents, each with a different skill set while homogenous activities will require that the group determine which of two similar or identical tasks is of higher priority and should be completed first, if the resources are not available to accomplish both simultaneously. This is very similar to the Earth observing satellite problem, as there will be multiple targets that require a range of sensors for adequate observations. Some targets will need the same sensor types to observe them, requiring that the group coordinate its activities to allocate the scarce resource of satellites with the needed sensor type. Heterogeneous targets will mean that while some satellites are available, they are not appropriate for the target that must be observed.

Ordering dependency and absence of ordering dependency

Ordering dependencies place a constraint on what order a set of tasks must be completed in. When no ordering dependencies are present, the set of tasks may be performed in any order. The presence of ordering dependencies may change throughout a problem and may only be applicable to different sub-sets of tasks, rather than to all tasks in the mission. The tasks that are necessary for the waste movement problem display both ordering dependencies and a lack of ordering dependency. Ordering dependencies include the requirement that the holding area door must be opened before the waste container may be placed into the holding area (meaning the door must be opened before all pushing tasks are completed) and the waste container may only be pushed a limited amount on one end before it is pushed on the opposite end. Non-ordering dependencies in the waste movement problem include the allowance of either end of the container to be pushed arbitrarily if the container is even and the flexibility of opening the holding area door during any point in the mission (as long as it is opened before all pushing tasks are completed).

Ordering dependencies and no ordering dependencies are also apparent in the Earth observing satellite problem. Depending on the requirement of the mission, some targets may require a certain viewing schedule between satellites with different types of sensors. For example, some targets may require observations be spaced out over a given amount of time, requiring an observation plan that will dictate a specific ordering of the over flights for each satellite, while some targets will not have this constraint imposed on the satellites, meaning that they have no ordering dependencies.

Agents can detect own actions

The assumption was made in both problems that each agent would have the capability to detect its own actions and to be able to interpret those actions as either successful or unsuccessful. This assumptions means that each agent can determine if the tasks that it has been assigned to complete are being properly performed, and if not, to relay that information to the rest of the group so that a different agent can begin that task.

Agents know about group activities only through communications

The assumption was made for both problems that agents can only determine the status of tasks through communications with other agents as opposed to passively "watching" the performance of other agents as they work on a task or on measuring the status of the task themselves. This assumption means that each agent is dependent on the group for mission status information, so that the agent can compile the feedback and determine for itself what is the status of the overall mission. This assumption also will result in a delay time if an agent fails at its task between knowledge of the failure and implementation of any solution. This is because other agents are not able to determine that the agent is failing while performing the task directly, but must rely on feedback from the agent performing the task.

No group optimization

Both problems have either no provisions or limited provisions for determining what group plan is optimal for completing the mission. In the waste movement problem, there is no optimization procedure that is used, though some limited tasking of agents towards an optimal solution is achieved through the implementation of a simple model based predictor in the mission. In the Earth observing satellite problem, once the EPOS 1.0 optimizer is used, no further optimization is used when determining what satellite should be reassigned to take over the tasks of any satellite that has experienced a failure. Though, like the waste management problem, some limited tasking of agents towards an optimal solution is achieved by using past optimal models to influence satellite reassignment.

Inclusion of a model based predictor to improve performance

A simple model based predictor was developed and integrated with the ALLIANCE behavioral planning algorithms to create a reaction planning algorithm utlized in both the waste movement and Earth observing satellite problems. This predictor had the goal of using a limited amount of available information to help influence the task assignments of the agents so that the group could accomplish the mission more efficiently than if no knowledge of the system was available. The model based predictor was added in only to increase performance and was not necessary to make the reaction based planner work in either of the problems.

Minimizing/Maximizing of objectives

The metric developed to measure the success of both problems was time. In the waste management problem, a highly successful mission minimizes the amount of time it takes to complete all the tasks, while in a highly successful Earth observing satellite problem a high successful mission will provide the maximum viewing time of all the targets possible.

Critical Differences

The following is a list and explanation of functional differences between the two problems.

Continuous tasks vs. Discrete tasks

The nature of the tasks that compose the two problems are relatively different. The major type of task in the waste movement problem is pushing the waste container. This task is of a continuous nature and takes a certain amount of time to successfully complete. During that time, the performance of the agent can be determined and if it is not satisfactory, a different agent can take over the task.

This is fundamentally different in the Earth observing satellite problem where the tasks are performing observations on targets where the time to complete the task may only last seconds or minutes. Here, as task is defined as each time a satellite views a target. While satellites typically perform orbital maneuvers in EPOS 1.0 that allow multiple observations to occur, one single observation is defined as a task. Multiple observations are defined as multiple tasks. If an agent is not able to accomplish the task, there is no time for a different agent to come and complete the task successfully. In the event of a failure, that viewing opportunity will be lost. The best solution that can occur is that a different agent will be assigned to observe a target and that no further viewing opportunities will be missed. The short time period that is available to successfully complete a task in the Earth observing satellite problem means that other group members will have a low patience level to allow an agent failing once at a task to try and correct the problem itself before the other agents intervene.

Physical path constraints

In the waste management problem there are effectively no physical constraints on the movement of the agents. It is assumed that the agents must stay on the ground, but are essentially free to travel anywhere else. Because there is no explicit path that they must follow, most traveling done by agents in the waste movement problem was not even modeled. In the Earth observing satellite problem there is a strong constraint on the movement of each agent. As each agent in this problem is a satellite, the agents are all constrained to move according to the laws of physics. This was accounted for in the problem by using the orbits propagated with the two-body problem in orbital mechanics. Any movement outside of these coasting orbits required the satellite to perform an orbital burn in a Hohmann transfer. Practical constraints limit the number and size of burns that are possible. This constraint on the movement of the satellites meant that a limited range of actions where possible to implement. Additionally, because the satellites are in orbit and the targets are fixed to the surface of the Earth, the targets will continuously be moving with respect to the satellites throughout the mission, because of the rotation of the Earth.

Fuel constraints

Fuel constraints will limit the amount of movement or reassignment that agents are capable of performing. In the waste movement problem, agents have no fuel limit and can move about freely. In the Earth observing satellite problem, each agent has only a limited amount of fuel that can be utilized and can therefore only perform a limited number of observations of a target. This means that if an agent fails, there is only a limited supply of fuel available to the group to try and compensate for the failure.

Centralized knowledge availability

The lack of centralized knowledge means that the group can operate in a decentralized manner, allowing each agent in the group to determine a course of action and implement the action without coordinated planning from the rest of the group. There is a complete lack of centralized knowledge in the waste management problem, but the Earth observing satellite problem has a limited store of centralized knowledge available to each agent. The knowledge base that is present in the satellite problem is generated from using the optimal planner before the outset of the mission. The optimal planner produces an observation and maneuver plan for each satellite that is optimal with respect to a developed utility function for the mission time horizon. These plans are generated even if the satellite is not actively used as part of the group. In the event of a failure by one or more members in the group, the remainder of the satellites can still compare their current situation with the old optimal plan generated, but never used. This gives the satellite some amount of centralized knowledge, but it is out of date.

4.3 Simulation Overview

The following section provides a discussion of how the simulation was setup and an overview of several of the scenarios that were examined in the simulation. An explanation of what was being studied in each scenario is also provided. Results for these scenarios and how the results are applicable to the Earth satellite problem are found in the following sections. These scenarios will be revisited later in the chapter to determine what effect the addition of a learning module will have on the performance of the group in accomplishing the mission.

Mission Development

Two different missions were used to aid in the development of the ALLIANCE behavioral planner in the waste movement problem that would later be applicable to the Earth observing satellite problem. The two missions are closely related, with the first mission being in effect a simplification of the second mission. Simplifications include a limited number of homogenous tasks (the door to the holding area is not present in the first setup), a smaller number of agents (three agents in the first setup compared to four in the second) and a constant terrain profile (the second setup has a variable terrain profile). A graphical representation of both simulation setups is included below, as Figs. 4.2 and 4.3.

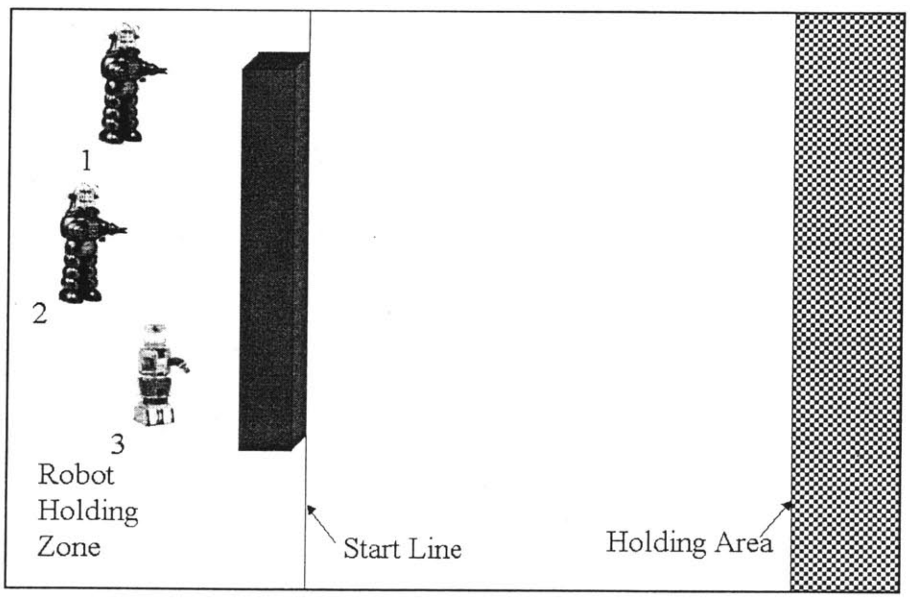


Fig. 4.2. Graphical representation of first simulation setup.

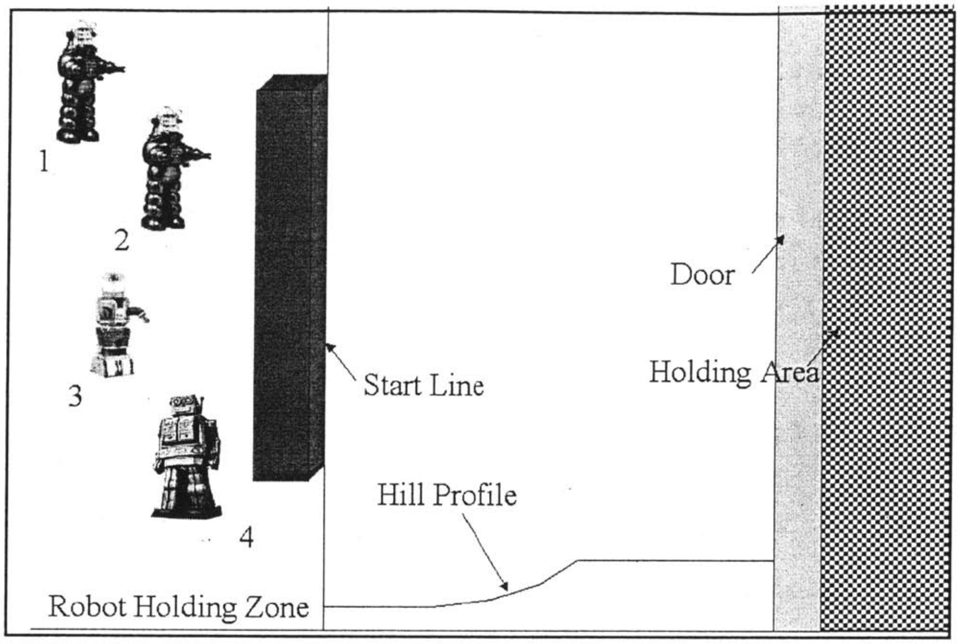


Fig. 4.3. Graphical representation of second simulation setup.

Scenario Development

From these two missions several scenarios were created. The scenarios were created to examine several parameters in the ALLIANCE and L-ALLIANCE algorithms. These include effects of failures on mission performance, parameter choices for each agent and effects of the learning module on mission performance. A complete listing of all studied parameters is provided below in Table 4.1.

Separate scenarios were developed to examine the effect that each of the studied parameters has on mission performance. The corresponding mission that these scenarios were developed for is also presented.

The output that is generated for each of the developed scenarios is presented below in the following section, along with a corresponding explanation. A short discussion is also provided on how these scenarios impact either the application of the ALLIANCE algorithms or the ALLIANCE algorithm's applicability to the Earth observing satellite problem.

Table 4.1. Complete listing of all scenarios and corresponding studied parameters.

Mission	Scenario	Parameter
Mission 1	Scenario 1	3 heterogeneous agents 1 total failure efficient use of agents
Mission 1	Scenario 2	3 heterogeneous agents 1 total failure 1 partial failure poor communication scheduling
Mission 2	Scenario 4	4 heterogeneous agents no failures learning module included model based predictor included initial poor impatience parameter setting

ALLIANCE Output

The following section provides an overview of the output that is generated using the ALLIANCE algorithms to plan and execute the various scenarios developed for the waste movement problem. This output will consist of motivation history charts for all scenarios, communications history charts for select scenarios and graphical representations of agent task performance for select portions of some scenarios. Explanation of all presented output is also provided along with discussion. How the output relates to modifying the ALLIANCE algorithms for use in the Earth observing satellite problem is presented in the following section.

Scenario 1 Output

Scenario 1 demonstrates how the ALLIANCE algorithms should be used. The scenario subjects the group to a failure that completely disables one of the group members and the remainder of the group must compensate to perform the mission. Parameters that affect the group's performance, such as rate of communications between agents and length of time that agents will allow other group members to work on a task before becoming impatience, are properly set so that the group works in an efficient manner. The motivational history of the group for the entire mission is presented below in Fig. 5.4. Some graphical representations of the group performing the tasks are also presented in the following series of figures. The correlation between these graphical representations and the motivational levels presented in Fig 4.4 will be elaborated on.

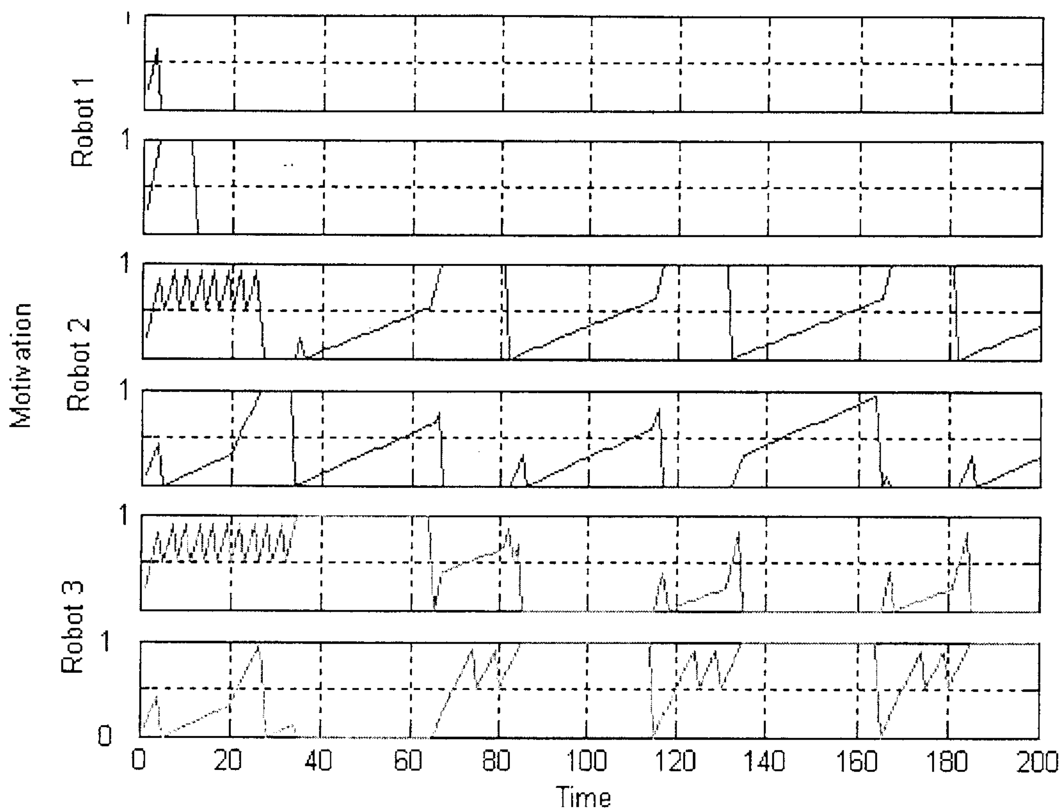


Fig 4.4. Motivation history of group in scenario 1.

The motivation history chart presented above can be briefly explained as follows. As there are three agents in the group and two tasks that need to be completed, there are six motivation histories that are presented, where each motivation history corresponds to the motivation that each agent has for each task. The first two motivation histories are for agent 1 on task 1 and task 2. The motivations of agents 2 and 3 are similarly presented. The x-axis presents the time that has elapsed since the mission began. When the motivation level reaches 1, then an agent will begin a task, unless a different agent is satisfactorily working on the same task. The motivation history presented in Fig. 4.4 is re-presented below in Fig. 4.5 with specific portions called out.

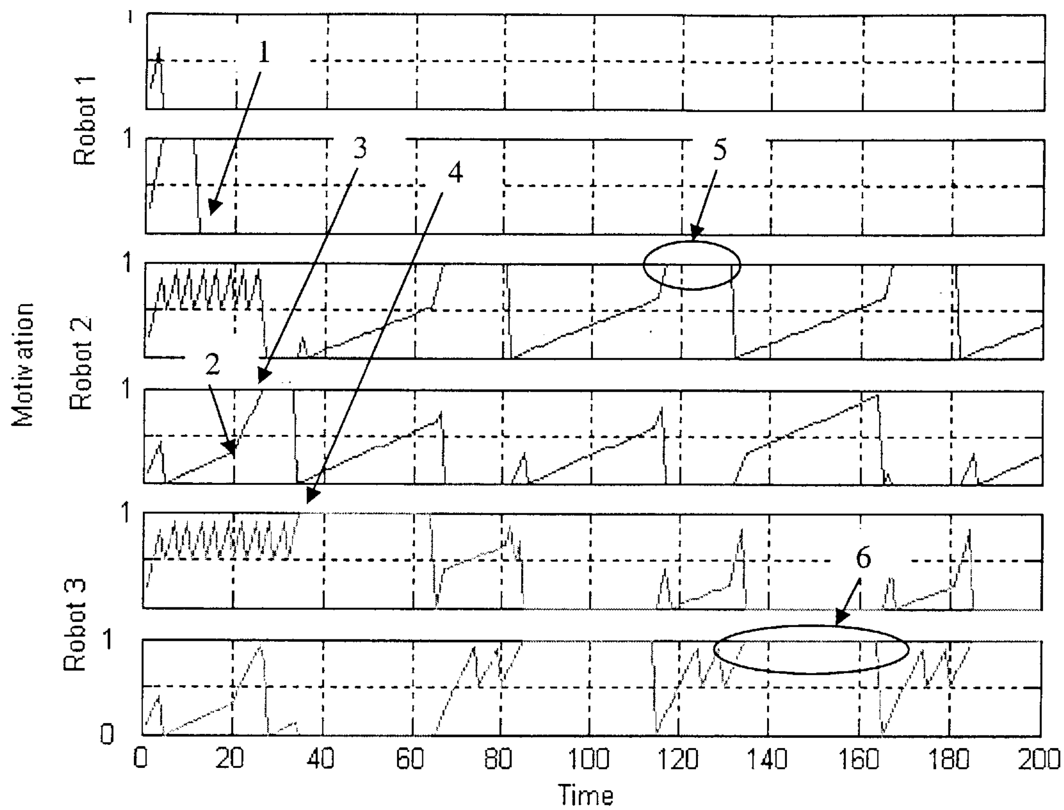
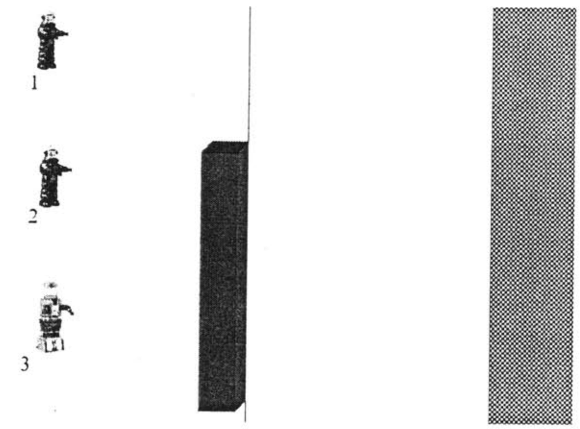
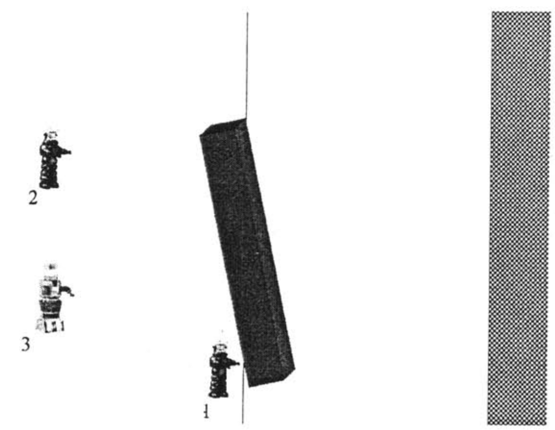


Fig. 4.5. Motivation history re-presented with callouts.

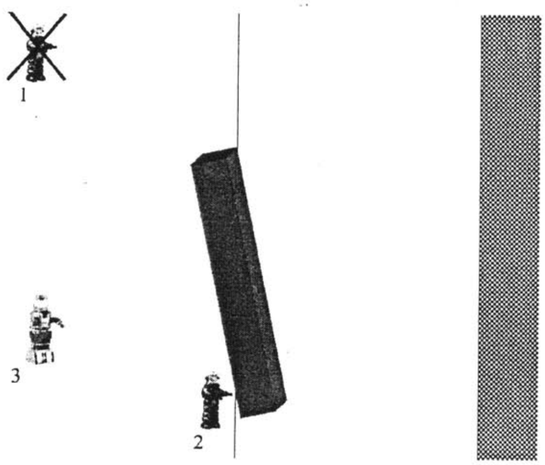
The mission begins with agent 1 being the first agent to be assigned a task. Part way through the task, agent 1 experiences a complete failure and is unable to either complete the task or alert the rest of the group about its failure. The failure of agent 1 is called out by Callout 1, which shows agents 1's motivation drop to zero. Because agent 1 is unable to communicate with the rest of the group due to its failure, after a pre-specified length of time has past without communications, the rest of the group's impatience begins to increase at a faster rate. This is shown for agent 2 at Callout 2. Eventually, agent 2's motivation passes a threshold, shown in Callout 3, and it completes the task that agent 1 started. This sequence is graphically displayed below in Fig. 4.6.



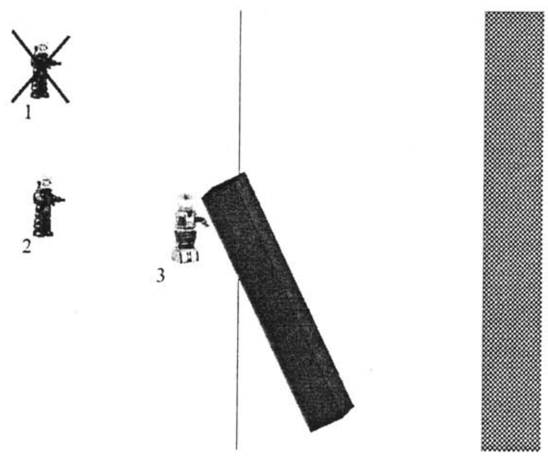
Time = 0 (pre-mission start)



Time = 7



Time = 25 (Callout 3)



Time = 35 (Callout 4)

Fig. 4.6. Graphical representation of mission.

The mission is efficiently completed because the agents are able to work together to complete the mission in the shortest time possible. They do this by starting each task just as the previous task is completed, as shown in Callout 4, and allowing each agent to work on a task without interruption, if the agent is accomplishing the task.

The group is composed of heterogeneous agents, where each agent is able to move an end of the container across the floor at different rates. Because of this heterogeneity, it takes each agent a different amount of time to accomplish a task successfully, as shown in Callouts 5 and 6.

Scenario 2 Output

Scenario 2 shows an example of a situation when the group is not able to work together as efficiently as desired. This can happen when all the parameters associated with the agents are not set properly with respect to the assigned mission. Values for each of the parameters of interest are set through experience with previous missions and a priori knowledge of the mission and agent performance. If any of these change or if there is not good knowledge of the mission before it begins, this effects the values assigned to the parameters which will affect mission performance. A detailed description of parameters that must be set prior to the mission is presented in Chapter 3. Scenario 2 shows an example of inefficient group behavior caused by an improper setting of the communication parameters with respect to the mission. This means that the agents are designed to communicate their status at the start and end of a task and at a set interval during a task. There is imperfect knowledge of the mission and as a result the length of time that each agent takes to complete a task is underestimated. This results in the agents not communicating often enough and other agents mistakenly believing that a failure has occurred. The resulting group behavior is one where the agents do not allow each other to complete a task but instead continuously start and stop a task as other agents take over. The motivation history for this scenario is shown below in Fig. 4.7.

Fig. 4.7 displays the same motivation histories as were previously explained for Fig. 4.4. In addition, the communication history for each agent is also displayed. The communication history is presented just below each agent's motivation history. Spikes that appear in the communications history show when an agent communicates with the rest of the group. Communications are defined as both transmitting the agent's status and uplinking the group's status. This occurs whenever an agent starts or stops a task or at predefined time intervals.

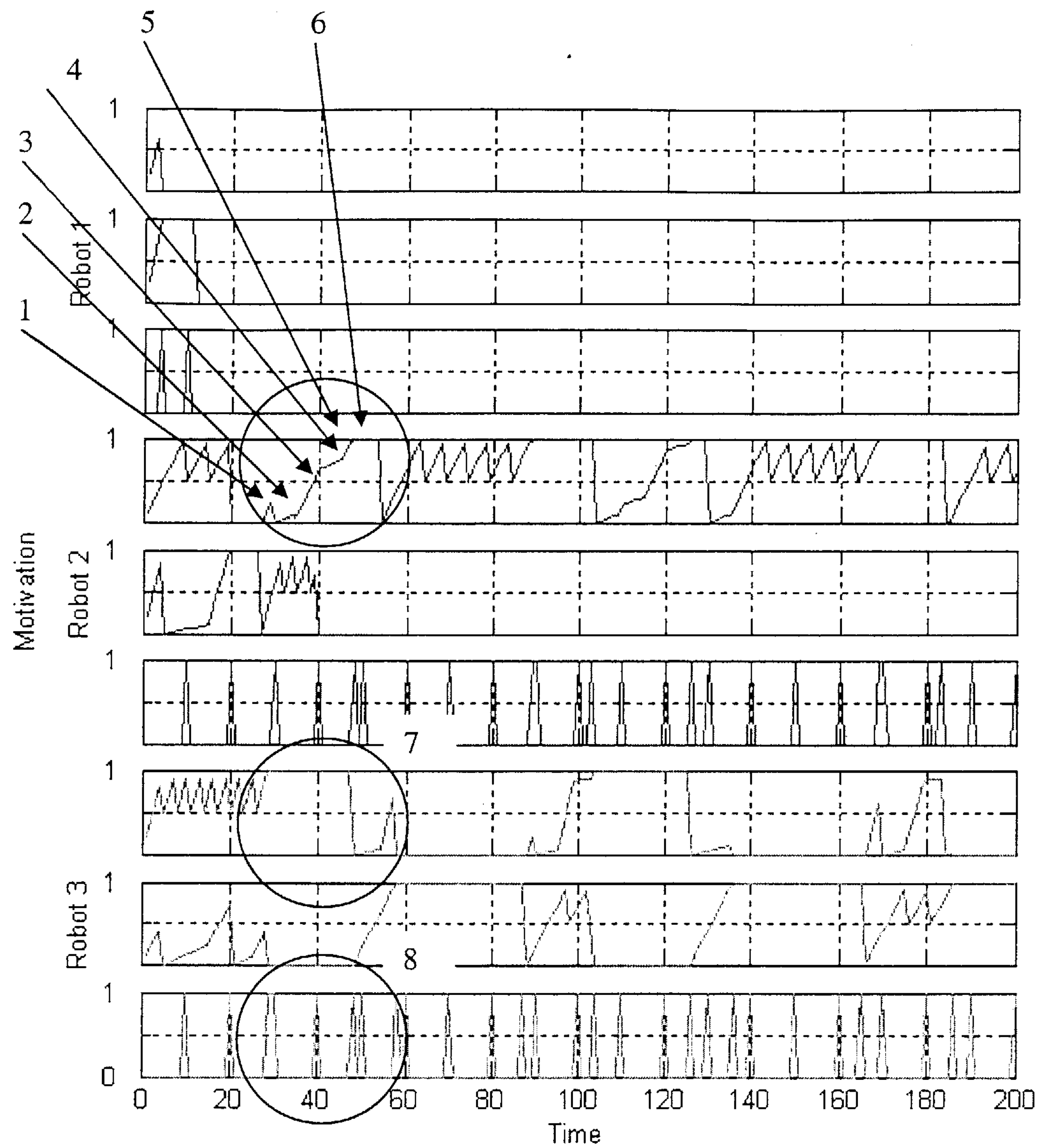


Fig. 4.7. Motivation and communication history during scenario 2.

As in the previous scenario, agent 1 begins a task and experiences a failure during the task. In this scenario agent 2 also experiences a failure, but it is not complete. The failure will not allow agent 2 to work on task 2, though agent 2 is allowed to continue work on task 1. Instead of agent 2 and agent 3 efficiently taking turns completing the two tasks, as was shown in the previous scenario, the two agent continuously interrupt one another while completing tasks, due to a mis-setting of the communications parameter for the mission.

Both agent 2 and 3 begin to grow impatient to complete task 1 around time 30 (Callout 1 and 7). Agent 3 reaches the motivation threshold first (Callout 7) and agent 2's motivation drops to zero (Callout 2). Because agent 2 was expecting agent 3 to complete the task sooner than it is able to, it expects a communication from agent 3, which it does not receive. This causes its impatience to increase at a faster rate (Callout 3). The rate of impatience is slowed in agent 2 (Callout 4) when a preset communications is received at time 40 from agent 3 (Callout 8). However, the same problem occurs again and agent 2's impatience again increases (Callout 5) until it reaches a threshold and, believing agent 3 has failed at its task, it begins working on task 1 (Callout 6). As a result it does not take agent 2 long to complete task 1.

This results in an inefficient use of group resources, as agents that are functioning properly are not allowed to complete their tasks before being interrupted. This occurs when the parameters have been improperly set. Reasons for this can include simple mistakes in assigning the parameters, lack of a priori knowledge about the mission, changing performance capabilities among the group or changing requirements for the mission tasks. The criticality of this type of sub-optimal performance depends on the specific mission and the resources available to the agents. If travel costs for the agent are very expensive, such as the case in the Earth observing satellite problem, then this level of sub-optimality is extremely critical. Likewise, this may also be critical if there is a loss in time where the task is being accomplished during the switching of agents. In the Earth observing satellite problem, if satellites were to constantly replace one another in this inefficient a manner, observation time may be lost while the satellites continuously keep changing their orbits to attempt and view new targets. Because of these two reasons, it is important in the Earth observing satellite problem to possess a higher degree of optimality in agent satellite assignment than is observed in this scenario.

The next scenario incorporates a learning algorithm so that the group can adjust its parameters if any of these situations occur.

Scenario 3 Output

Scenario 3 utilizes the learning algorithm so that the problems encountered in scenario 2 can be avoided. Scenario 3 does not have any failures that the group encounters, though it does have an initial ineffective setting of parameters, similar to that found in scenario 2. The parameters that are incorrectly set in scenario 3 deal with the ability of each agent to accomplish the given tasks. Before the mission begins, each agent's abilities are known from previous missions or from their design and this knowledge is supplied to each agent, which then affects which agent will be assigned tasks. In this scenario agent 3 is believed to be the best performer for tasks 1 and 2 and is thus repeatedly assigned to perform them. However, agent 3 suffers degradation in its ability to perform the tasks and as a result it takes much longer than expected to complete the tasks. The other group members learn about this degradation in agent 3's ability and as a result update their databases on the capabilities of each agent. Finally, agent 3 is no longer the primary agent to complete the tasks because of its degradation in performance. This is shown below in the motivation history presented in Fig. 4.8.

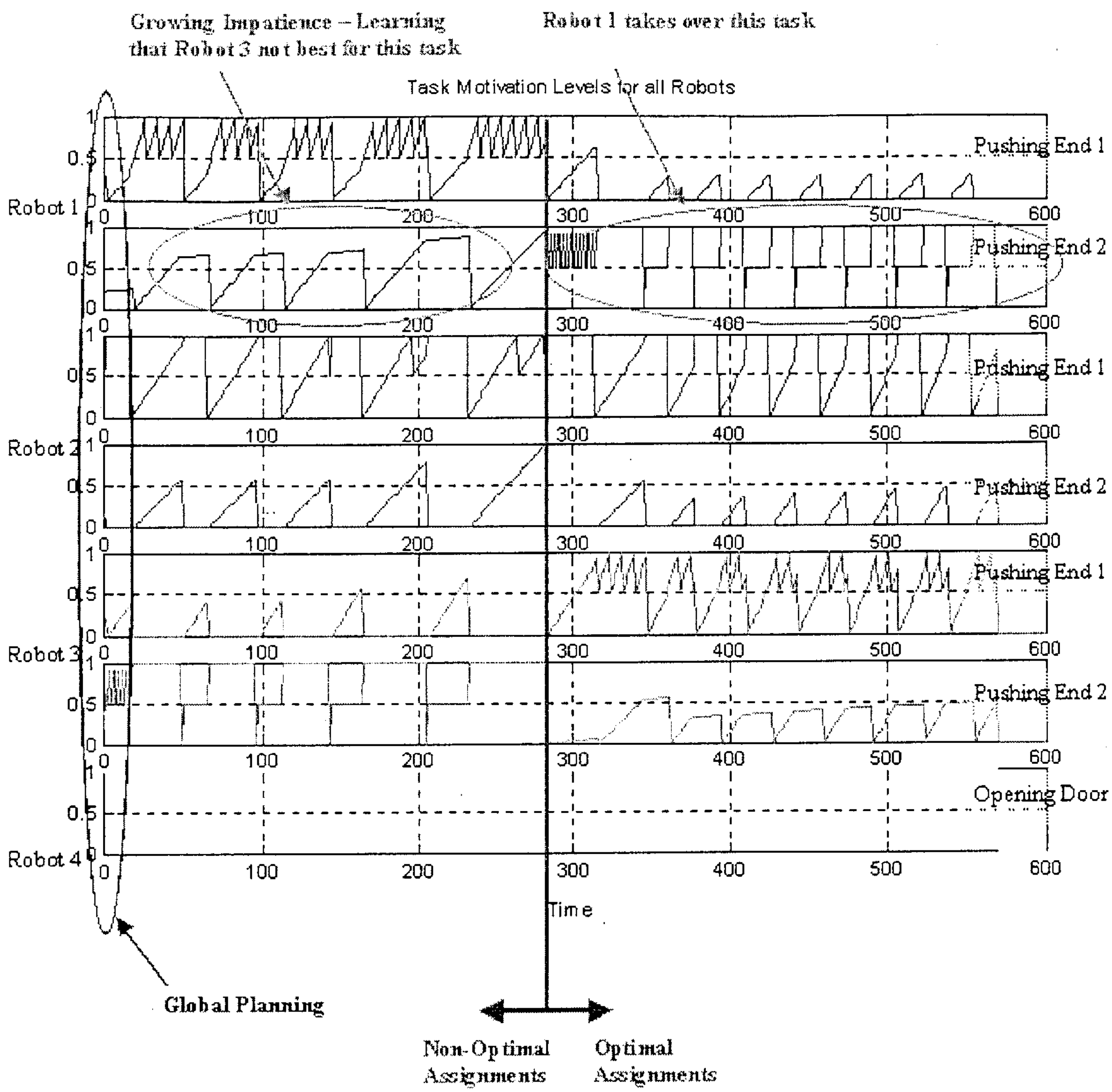


Fig. 4.8. Motivation history for scenario 3.

Also included in this scenario is a capability to analyze all the tasks in the mission and complete an initial assignment of each of the agents to a specific task so that the entire mission can be effectively accomplished.

Modifications to ALLIANCE and L-ALLIANCE Algorithms

Two major modifications were made to the ALLIANCE/L-ALLIANCE algorithms in the waste movement problem. These modifications were deemed necessary after working with the original algorithms and determining that improvements could be made in coordinating the group to more efficiently complete the mission. These two modifications examine the occurrence of dynamic tasks and implementing global planning for the group. A description of the modifications made to the algorithms is presented below.

Dynamic Tasks

Dynamic tasks posed a problem for the efficient application of the ALLIANCE/L-ALLIANCE algorithms. Dynamic tasks are defined for the waste movement problem as tasks that change in difficulty throughout the mission. Task difficulty is measured in the ALLIANCE architecture as time required to complete a task, meaning that tasks that take a longer time to complete are more difficult. While the ALLIANCE architecture is designed to accommodate a mission with variable task difficulty and is able to handle dynamic tasks, it was designed to primarily handle tasks that possess the same difficulty throughout the mission.

The problem that dynamic tasks pose is primarily to the L-ALLIANCE learning algorithms. The learning algorithms are designed to update each agent's knowledge of all other agent's performance capabilities based on an agent's performance throughout the mission. An agent's performance capability is a function of how quickly it can complete a given task. The shorter the time for completion, the better the performance capability. Once an agent completes a task, its performance capability is updated based on its latest performance. This updated information is broadcast throughout the group so that the entire group then updates its information on how well an agent can perform a task. This updated information will then be used to readjust all the parameters that affect an agent's impatience and acquiescence behaviors. For example, if an agent performs a task better than expected, all agents in the future will expect a slightly better performance out of that agent when performing that task.

When a task takes longer to complete than expected, this poses a problem for the efficient utilization of group resources. When a task takes longer than expected to complete, the

L-ALLIANCE algorithms are designed to assume that the agent performing the task is not functioning as well as it should. While this may be the cause for the increase in the time required to complete the task, another reason may be that the task itself has become more difficult during the course of the mission and that the same performance from the agent will simply mean that a longer amount of time is required to complete the task. Unfortunately, what can occur is that the rest of the group believes that the agent working on the now more difficult task is not functioning well enough and will relieve that agent from the task and a different agent will begin the task. This unnecessary re-tasking of agents is inefficient as the group wastes resources and time continuously by traveling to and from the task site.

A solution to this problem is attempted through the introduction of a model-based predictor to the basic ALLIANCE architecture. The model-based predictor is implemented to address the subset of dynamic tasks that are known to change in difficulty throughout the mission, while the subset of dynamic tasks that include changes that are unknown before the mission begins are not addressed in this research. The second simulation setup addresses this type of dynamic task. This is represented in the second simulation setup by requiring that the waste container be pushed up a variable sloped incline. It is assumed that the terrain that must be traversed is known before the mission begins and that the time required for each agent in the group to push the container up the hill a set distance can be determined based on the agent's nominal performance and the slope of the hill.

The predictions on future performance supplied by the model-based predictor were integrated in with the agent's previous use of past performance so that impatience levels of the rest of the group would be based on both, one, a prediction of how well an agent would perform a task based on its current abilities and the difficulty of the task, and two, how well the agent has been performing a certain task. This affects all the parameters that are determined by the $task\_time_i$ parameter, as described in Chapter 3. The value assigned to $task\_time_i$ was calculated in several ways to represent the precision of knowledge available about the tasks before the mission begins. Three different schemes were used to represent the knowledge of the task difficulty, including certainty of knowledge, and two degrees of uncertainty, these being a low and high uncertainty level. The $task\_time_i$ was calculated in the following manner for each of these knowledge states.

$$task\_time_i(t) = \text{predicted performance} \tag{4.1}$$

$$task\_time_i(t) = \text{average(predicted performance and past agent performance)} \tag{4.2}$$

$$task\_time_i(t) = \text{average(predicted performance and past agent performance)} + \text{one standard deviation of average} \tag{4.3}$$

Global Planning

Global planning was implemented within the ALLIANCE architecture to help coordinate group task distribution more efficiently. One of the defining characteristics of the ALLIANCE architecture is that it is designed to allow the group to operate in a very decentralized manner to accomplish the mission. This is done by each agent being assigned to a task as it become available, based on the agent's ability to complete a task. There are no provisions in the ALLIANCE architecture to assign tasks to agents based on the desire to increase the efficiency of the overall mission. Even though the most capable performing agent that is available is assigned to each task as it becomes necessary to complete, this can result in an overall decrease in mission efficiency. To increase mission efficiency, it may be necessary to assign less capable agents to some tasks. An example of when this is necessary to increase group efficiency is modeled in the second simulation setup. This occurs when assigning the heterogeneous agents the tasks of either opening the door to the holding area or moving the container across the room.

In the second simulation setup, there are four agents, two of which are homogenous and have the best performance capabilities in the group for all the tasks and the other two which have different levels of performance capabilities form each other and from the previous two agents. During the mission, one of the best performing agents experiences a failure, leaving three heterogeneous agents to accomplish the tasks of opening the door to the holding area and moving the container across the room and into the holding area. The task of moving the container across the room is subdivided into many subtasks of repeatedly pushing on alternating ends of the container until it is moved across the room, while the task of opening the door consists of one long task of traveling across the room, opening the door and traveling back.

If the ALLIANCE architecture is used, the best performing agent will begin the task that takes the longest time to complete, which in this scenario means that the best performing agent will be tasked to travel across the room and open the door, leaving the other two agents the task of moving the container across the room. The result is that the mission takes longer to accomplish than is possible. The "optimal" assignment of tasks in this scenario is for the agent with the lowest performance capability, and hence the slowest, to travel to the far end of the room and open the door. While doing this the better performing agents are left with the task of moving the container across the room, which takes a longer time than just moving across the room. The result is that the better performing agents are able to move the container across the room in almost the same time that the lowest performing agent is able to open the door. The assignment of agents in this manner would not have been possible without looking at all the tasks and agent performances at the outset of the mission and then determining what was the assignment of agents to tasks that would best accomplish the mission.

This is accomplished by examining all the combinations of agent task assignments that are possible and, based on the performance predicted in the model-based predictor, assigning the agents to the tasks that provide the lowest total mission completion time. This assignment is executed by weighting the impatience levels of each agent for the appropriate task at the outset of the mission. The weighted impatience level results in the agents becoming motivated to start the tasks determined to have a high mission efficiency. After the mission begins, agent impatience levels are calculated according to the ALLIANCE and L-ALLIANCE algorithms, as explained in Chapter 3. The impatience level at the outset of the mission is calculated in the following manner.

$$\delta\_fast_{ij}(0) = weight * \delta\_fast_{ij}(pre\text{-}mission) \quad (4.4)$$

where i is the agent best assigned to task j to achieve an efficient mission completion.

This combinatorial approach to finding a more efficient assignment is only practical when the number of combinations is reasonable and the knowledge of the mission and agent performance is known with reasonable accuracy.

ALLIANCE and L-ALLIANCE Applicability to Earth Observing Problem

Applying the ALLIANCE and L-ALLIANCE algorithms to the waste container movement problem helped in the understanding of how the algorithms functioned and how they could then be applied to the Earth observing satellite problem. While the Earth observing satellite problem will be described in more detail in the following chapter, this section discusses some of the critical applicability issues when using the ALLIANCE architecture to provide operational support to the group of satellites during unexpected events.

The most critical issue that poses a challenge for applying the ALLIANCE architecture to the Earth observing satellite problem is the manner in which tasks are defined, including a time critical component. ALLIANCE and L-ALLIANCE implicitly assume that a mission will be formed of several tasks, each taking a certain amount of time to accomplish. The time that it takes to accomplish the tasks is assumed to be a substantial fraction of the total mission time and the time that it takes to accomplish any one task is not inconsequential. ALLIANCE is designed to allow other agents to essentially monitor the progress an agent is making at completing a task while it is working on the task. If the agent is not performing that task well, after a set amount of time other agents will become impatient and take over the task from the first agent. It is assumed that the task will be available in the same state until the next agent arrives to begin work and that this lapse in time between agents working is inconsequential. This is not the case for the Earth observing satellite problem.

In the Earth observing satellite problem the mission consists of viewing one or more targets several times over the course of the mission. The amount of time that is required to view each target is very short when compared to the mission. The target observation time is on the order of seconds or minutes while the mission is on the order of days or weeks. During the rest of the mission while the satellite is not observing a target it is traveling to the next target for a future observation opportunity. The short time that is available to observe a target is a critical problem for ALLIANCE, which assumes that there is time while an agent is performing a task to determine if it is performing well. Additionally, in the Earth observing satellite problem if a satellite is found to be not capable of viewing a target once it is overflying that target, there is not enough time for another satellite to recognize the problem and travel to the target so that it can be observed at the same time that the first satellite was going to observe it. At best, the new satellite can observe the target at a later point in the mission, but the original observation can never be recovered. The short time duration of the tasks, the long travel times to the target and the inability to perform the same task at a later date poses a challenge for the ALLIANCE architecture to be applied effectively to the Earth observing satellite problem. These problems are addressed in the next chapter.

Another application issue in the Earth observing satellite problem is the choice of metric to determine if the satellite is performing its tasks effectively. Both the ALLIANCE and L-ALLIANCE algorithms use time as a metric to judge the performance of an agent. If an agent can accomplish a task in a short amount of time, it is judged to be performing effectively. L-ALLIANCE's learning algorithm compares the time it takes an agent to complete a certain task each time it works on that task, to update the groups knowledge on how well an agent can perform a task. The use of time as a metric in the Earth observing satellite problem to measure performance may not be the best choice. Here, time to complete a task is not necessarily an indicator of a satellite's performance, but is more a function of the orbital mechanics governing the satellite's motion. Additionally, each satellite is trying to maximize the amount of time that it can see the target. This means that each satellite wants to maximize the amount of time it takes to complete the mission, not minimize it as is assumed in ALLIANCE and L-ALLIANCE. A better performance indicator for the Earth observing satellite problem may be an efficiency ratio that relates the time a satellite can observe a target to the amount of fuel or time that it takes the satellite to reach the target. Other performance indicators could also be employed, which either judges an individual satellite's performance or its performance as a member of the group.

Because of these characteristics of the Earth observing satellite problem, the ALLIANCE algorithms were modified in several ways when applied to the Earth observing satellite problem. These modifications are discussed in the following chapter. Also, the L-ALLIANCE algorithms were not used in the Earth observing satellite problem. This is because the learning algorithms assume that past performance is a good indicator of an agent's future performance. For the Earth observing satellite problem, future performance seems to be in large part a function of the orbital mechanics that the satellite is subjected to rather than a performance issue. If a different metric besides time is used to judge performance, L-ALLIANCE maybe more useful to the Earth satellite problem. This is an issue that is saved for additional studies. However, it is recommended that some type of learning should be included, either prior to the start of the mission or during the mission, to fine tune the various parameters in ALLIANCE.

Chapter 5

Integration of Optimal and Reaction Planning for Earth Observation

A reaction planning algorithm was integrated with the EPOS 1.0 pre-mission optimal planner, which creates a planner that provides robustness to the mission in the event of one or more satellite failures. This chapter provides an overview of the work completed in integrating the two types of planners together to form one integrated planner and its application to the Earth observing satellite problem through the development of a simulation environment. Included in this chapter is an explanation on how the integrated planner works, followed by modifications that were made to the EPOS 1.0 optimal planner and reaction planner to allow their integration. Also included is an overview of the simulation environment that was created to test the performance of the integrated planner and the scenarios that were created for the simulation environment. The chapter is concluded with a step-by-step sample run of one scenario that was developed to illustrate the mechanics of the integrated planner.

5.1 Overview of Integrated Planner

The integrated planner is comprised of two major sub-systems; the EPOS 1.0 pre-mission optimal planner and the reaction planner. These two sub-systems are integrated together to form a planner that is; one, capable of generating plans before the mission begins with the pre-mission optimal planner which optimizes the benefit obtained and the cost expended for the satellite group and, two, provides mission robustness with the reaction planner in the event of an unexpected failure once the mission has begun. These two types of planners are integrated together to provide functionality for both the need to efficiently allocate the scarce fuel resources of the satellite group and to adapt the plan that is generated for observing the targets in the event of an unforeseen failure during the course of the mission. Fig. 5.1 displays a flowchart that illustrates how the optimal planner and the reaction planner are integrated together. The following discussion provides an overview of how these two planners operate together.

The first four boxes displayed in Fig. 5.1 display the major steps involved in the EPOS 1.0 pre-mission optimal planner. All planning activities in these steps occur before the mission begins and the satellites have been given any assignments. The first box represents the user data inputs that must be given before any planning can be accomplished. Data that is necessary for the planner to function is summarized below in Table 5.1. This data represents information that is necessary to determine what target the satellite group will observe and constraints on how the observations are to be obtained, in terms of time, number of satellites available and the ability to perform orbital maneuvers.

Table 5.1. Summary of user inputs for EPOS 1.0 pre-mission planner.

User Inputs	Purpose
Identify target or targets of interest	Targets on which information will be gathered
Determine length of mission time horizon	Length of time that targets are of interest
Specify benefit metric	Importance of type of observation (examples, total length of observation time, number of observations, maximum gap time between observations, etc,)
Allow burns or restrict to coasting	Allows or disallows satellites in group to expend fuel to perform orbital maneuvers
Allocate the maximum number of satellites available for the mission	Maximum number of satellites that are allocated to viewing the target

The second box, labeled "optimizer", represents the optimal planner that was developed for EPOS 1.0. The optimal planner is engaged before the mission begins and is tasked with the goal of determining the optimal tradeoff between target viewing time and fuel that is expended through performing orbital maneuvers for each satellite in the group. The optimal planner is designed to increase the amount of viewing time that each satellite can obtain by allowing the satellites to perform orbital maneuvers. These orbital maneuvers place the satellite in an orbit that will increase the amount of time that the satellite can observe the target. As performing orbital maneuvers can expend a significant amount of fuel available to the satellite, and fuel is a scare resource, there is a limit on the number, size and types of orbital maneuvers that can be performed. The trade off between maximizing the viewing time, or benefit, that is obtained by allowing orbital maneuvers and minimizing the fuel expended, or cost, when performing orbital maneuvers is determined through the optimal planner. The mathematics of the optimizer were explained in detail in Chapter 3.

The third box represents the task of creating groups. This task is accomplished by the user, with the decision making process aided with relevant information that is provided by EPOS 1.0. The user creates a group of satellites based on a set of requirements or criteria that is important to a particular mission. There are potentially many sets of satellites that can be tasked to observe the desired target and the actual set of satellites that will be utilized is chosen by the user using the appropriate information supplied by EPOS 1.0. A detailed example of this procedure is provided later in this chapter.

The fourth box represents the implementation of the plan and the start of the mission. After the set of satellites are chosen from the previous set, a schedule is created for each satellite. This schedule includes a maneuver plan that details when each satellite is to perform an orbital maneuver and what type and size of maneuver to perform. The maneuver plans are uplinked to the individual satellites, represented by the box called "Information to and between satellites", and the mission then begins at a predetermined time. The actual implementation of the plan during the mission is represented by the large gray box called "Plan in Effect". Also created is an observation schedule that informs the user of the precise time that the target will be observed and by which satellite.

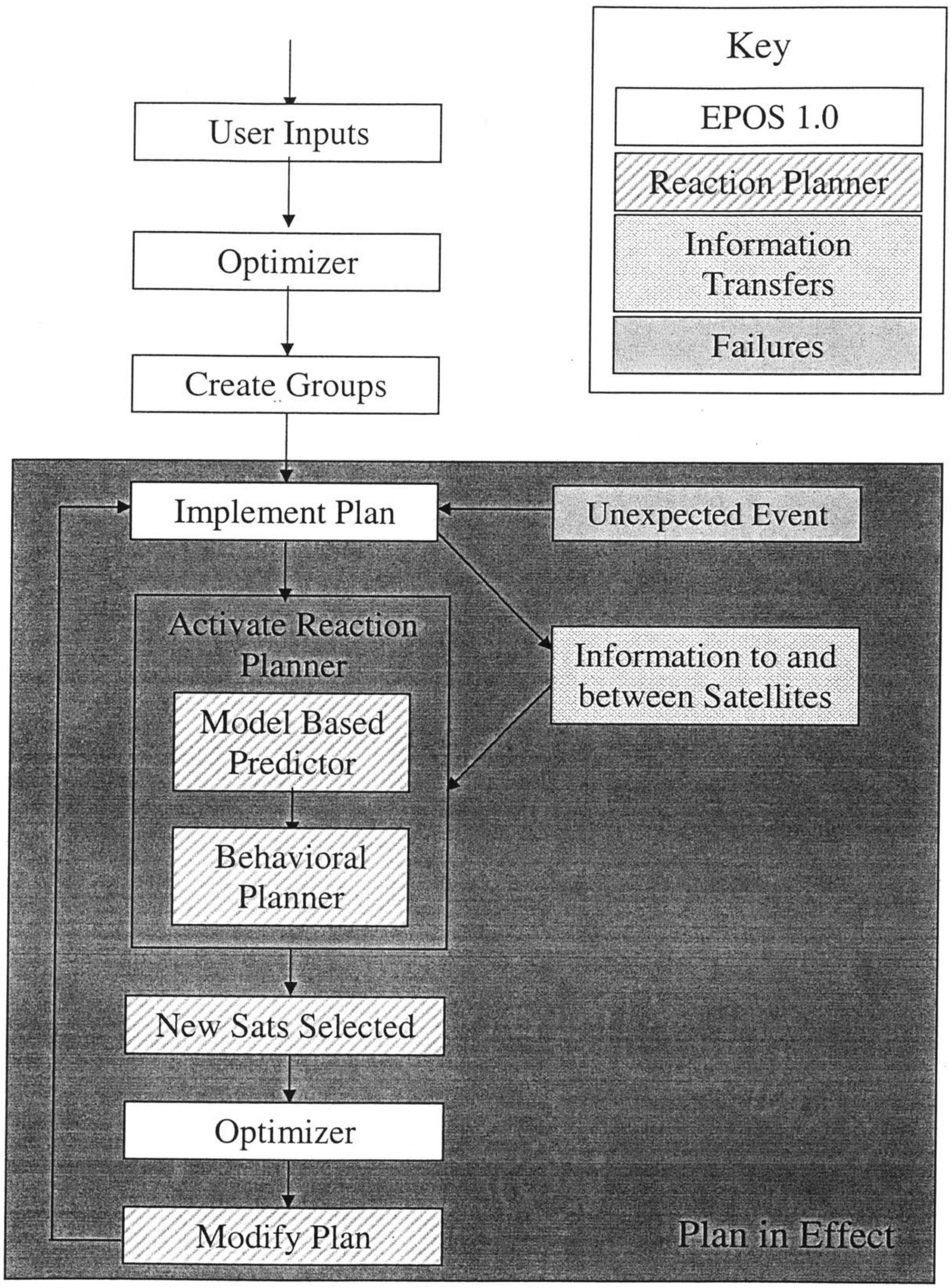


Fig. 5.1. Integrated planner flowchart.

Once the plan has been implemented, the mission is performed as determined by EPOS 1.0, unless an unexpected event occurs that could significantly affect mission performance. Unexpected events can take the form of satellite failures, changing observation requirements of targets or new opportunities. These unexpected events and some examples of each are listed in Table 5.2, below. For this thesis, only the first two types of unexpected events are examined.

Table 5.2. Unexpected events and examples.

Unexpected Events	Examples
Satellite failure	Complete or partial failure of one or more satellite subsystems – loss of observation equipment, power failure, lose of ability to perform orbital maneuvers, loss of communications. Failures of this type may be permanent or temporary.
Changing observation requirements of target	Satellites originally tasked to observe target may not be able to due to changing environment – satellite with visual sensor may not be able to see through cloud cover.
New opportunities	New targets of interest may appear after the initial plan has been implemented or new event may occur suddenly with original target that should be observed.

When an unexpected event occurs, the plan that is being implemented is damaged, as the observations that were planned can no longer all be performed. To regain some of the observations that were lost due to the unexpected event, a new plan must be formulated. This is accomplished using the reaction planner portion of the integrated planner. The reaction planner analyzes the unexpected events that have occurred and attempts to fix the plan. This procedure is illustrated in Fig. 5.1 and described below.

The reaction planner is activated after an unexpected event occurs that affects the original plan that is being implemented. This sequence of events is shown in the boxes labeled "Implement Plan", "Unexpected Event" and "Activate Reaction Planner".

Before the behavioral algorithms that are based on the ALLIANCE algorithms, as described in the previous chapter, are activated, the reaction planner engages the model based predictor to determine which satellites are the most appropriate to repair the damaged original plan. This analysis is accomplished using data generated when using the EPOS 1.0 pre-mission planner and is assumed available to each of the satellites at the outset of the mission. The end result of using the model based predictor is that each satellite that is applicable to be used in repairing the plan is weighted according to the benefit it can provide. Satellites that can produce more observation time are given a higher weighting and will be more likely to be selected when the behavioral planning algorithms are engaged. This procedure is represented by the box labeled "Model Based Predictor". Alternatively, if the data used in the model based predictor is not available or is not current enough, the model based predictor can be entirely bypassed and the behavioral planning algorithms can be directly engaged. This has the effect of weighting all satellites equally. The output of both approaches is that a set of new satellites is selected to repair the damaged original plan.

Once the satellites have been selected, the maneuver and observation plan must be generated. The maneuver plan determines what orbital maneuvers will be performed and when and the observation schedule will be a resulting schedule of when the satellite can view the target. This is accomplished by using the same optimal planning software as used in EPOS 1.0. The resulting plan for the replacement satellites has a balance of observation time and fuel expended while performing orbital maneuvers. This procedure is represented in the box called "Optimizer" that is located in the large gray "Plan in Effect" box. The result is a modified group plan that is then implemented.

5.2 Modifications to the Behavioral Planner

Several modifications were made to the ALLIANCE algorithms to address issues raised in the previous chapter concerning the application of ALLIANCE to the Earth observing satellite problem. These modifications were variable definitions of tasks, impatience weighting and variable definitions of task failure. These are discussed below.

Variable Task Definitions

One of the problems identified in applying the ALLIANCE algorithms to the Earth observing satellite problem is that the task of observing a target occurs in a relatively short time frame allowing no time for other satellites to become impatient and take over the task in the event of a failure. This problem was partially addressed by redefining a task as not only being the observation of a target, but also the orbital maneuvers that a satellite must perform to enable it to observe the target. By including the travel time in the task definition, the time-based component of ALLIANCE is again applicable for use. It is assumed that during the time it takes a satellite to travel to a target, self-diagnostics can be performed that increase the probability that a failure will be detected. Note that defining tasks in this manner does not completely eliminate this problem for two reasons. First, it is assumed that there will be times when either the self-diagnostic is not able to detect the failure and the failure is not identified until the satellite is ready to actually observe the target. The second possibility is that there is no time, capability or little desire by mission operations personnel to periodically utilize the satellites resources to perform self-diagnostics. In these instances there will be no time for other satellites to re-task themselves to observe the target in the time frame allotted to the initial satellite. Even if the satellite can perform a self-diagnostic during the time it is maneuvering in route to the target, if it identifies a failure shortly before it is supposed to observe the target there may still be no time for another satellite to be re-tasked to observe the target during the same time period as was originally scheduled. The best solution that can then be enacted is for the plan to be repaired or to "make up" the lost viewing time later in the mission.

Tasks were defined for the Earth observing problem as the actual act of observing a target and the time that is required to maneuver from its current position to a position capable of observing the target. The actual time required to perform the task varies as a function of when the target will next be in view, based on legal orbital maneuvers, as defined by EPOS 1.0 and explained in Chapter 3. A graphical representation of this is shown below in Fig. 5.2. This task definition is the nominal definition used to detect satellite failures.

This task definition is changed however to respond to environmental events that would prohibit the satellite from seeing the target, which in these simulations was modeled as cloud coverage. Whereas it was assumed that satellite failures would have a high probability of occurring with no warning, it was assumed that cloud coverage over a target could be predicted with some degree of accuracy a certain amount of time prior to the satellite actually arriving at the target. For this simulation, it was assumed that 2 hours prior to a satellite passing over a target that the possibility of cloud coverage obscuring the target from the satellite could be determined. If this was the case, then the rest of the group could immediately be notified and a satellite with a different sensor type impervious to cloud coverage could be tasked.

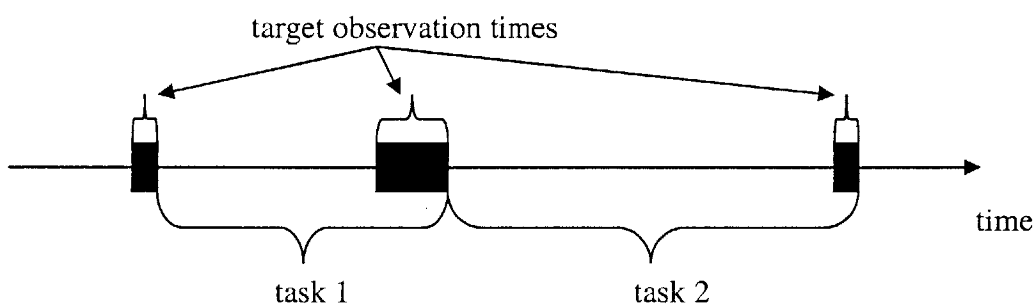


Fig 5.2. Sample task timeline.

Impatience Weighting

To increase the efficiency of the group a weighting method was developed that tied impatience growth levels to expected viewing opportunities. Expected future viewing opportunities were determined with the model-based predictor, which used previous data generated from EPOS 1.0 on satellite observation times. Satellites that could be expected to have a greater amount of observation time in the future would have an increased impatience for performing a task in the event of an unexpected event occurring. If no other activities are occurring, then it is highly likely that the satellite with the impatience that has experience the greatest weighting will be tasked for observing the target. In the event of multiple unexpected events occurring, each satellite is weighted according to its ability to perform at a certain task. As satellites become assigned to tasks and other tasks still need satellites to fulfill them, the satellite with the highest weighting may not be chosen. As a simple example, if two unexpected events occur around the same time in the mission, satellites in the group will be notified and their respective impatience will increase. If one of the satellites would be the best satellite to accomplish both tasks, it will have the highest weighting assigned to its impatience for both tasks. As the satellite cannot be assigned two tasks simultaneously, it will likely be assigned one task and the satellite with the second highest weighting will likely be assigned to the other task.

Weighting is determined by taking the expected observation time that each satellite will have from the current time to the end of the mission and determining what percent of the total time that is possible each satellite possesses. This percentage is then multiplied to the δ_fast variable, which influences the satellite's impatience growth level.

Variable Failure Definitions

Failure to accomplish a task is defined according to the capabilities of each satellite and the situation it encounters. Four types of satellites were modeled for the Earth observing satellite problem. The satellites in the group either had a visual or an infrared sensor, with the group being comprised of 50% of each type of satellite. Some of the satellites also had the ability to perform self-repairs if diagnosed with a failure, while the remainder of the group did not have this capability. Satellites with visual sensors could not see through clouds, while cloud coverage had no effect on satellites with infrared sensors. The group response to failures is determined by these attributes. If a satellite with no self-repair capability experiences a failure, the rest of the group immediately reacts to this and will re-task a comparable satellite, if available, to finish the mission in place of the failed satellite. If a satellite with self-repair capabilities experiences a failure, the rest of the group will slowly increase their impatience but wait a given amount of time for the satellite to repair itself. If it cannot, or cannot in a given amount of time, it is considered a failure and the remainder of the group acts as described above. In the event of cloud coverage, if clouds are expected to obscure a visual satellite two viewing opportunities in a row, a satellite with an infrared sensor, if available, will immediately be tasked to view the target. If cloud coverage is expected while an infrared satellite is viewing the target, this is not considered a failure, as cloud coverage has no effect on its operation.

This variable definition of satellite failure is critical given that the mission is composed of satellites with differing capabilities to repair failures and react to the environment. This variable definition extends the homogenous definition of failure that is utilized by ALLIANCE, which is if an agent cannot complete its task in a certain time frame, it is considered a failure.

5.3 Overview of Simulation Environment

The following section provides an overview of the simulation environment that was created to test the effectiveness of the integrated planner. This section covers the mission that is assigned, the satellite group that is tasked to complete the mission, the environment and different failures and events that are encountered.

Mission Profile

The mission that is to be accomplished is the observation of one or more targets comprising Earth based phenomenon of interest by a group of Earth observing satellites over a specified time horizon. Two goals that are in place during this mission are that, one, the satellites are to be used efficiently, in terms of fuel usage, and, two, the satellite group should be able to respond to unexpected events occurring after the mission has begun. The first goal is accomplished through the use of EPOS 1.0's optimal planner, which attempts to maximize the viewing time that is achievable with each satellite by performing orbital maneuvers while limiting the amount of fuel that is expended in performing this action. The second goal is achieved through the use of the reaction planner. The unexpected events of satellite failure and changing viewing requirements are met with the reaction planner.

During the course of the mission, two major types of activities must be completed. These activities are, one, performing observations for a limited time while the satellite over flies the target, and, two, traveling from the satellite's current position to a position that will over fly the target and permit an observation to occur. This second activity may entail the satellite to perform either an orbital maneuver or just simply coast. These two activities are combined to create a set of tasks that each satellite group must perform to successfully complete a mission. A task is defined here as comprising the movement of a satellite from a current position to a position capable of viewing the target and concludes with the over flight of a target. For most cases a task begins immediately after a satellite has viewed a target and is preparing to view the next target and ends after the end of the next target viewing. The task includes the time that is required to travel to the target. During this time an orbital maneuver may or may not be performed. Tasks were defined in this manner because a successful viewing of the target will require that the satellite has the ability to travel to the target.

A combination of two metrics is used to gauge success of the mission. The first metric is the amount of time that the satellite group can view the target. The second metric is the amount of fuel that is used in viewing the target. The effectiveness of the satellite group in accomplishing the mission can be measured by taking the ratio of the viewing time achieved with the fuel expended. This creates a ratio measuring the amount of fuel expended for every minute of viewing time achieved. Effective groups will be able to have a high ratio, or a long viewing time while a minimal amount of fuel is consumed. This set of metrics and measures can be applied to the integrated planner and its two sub-systems, the EPOS 1.0 optimal planner and the reaction planner.

Satellites

A group of 24 satellites was used to implement plans generated by the integrated planner. The group begins the mission in a Walker constellation of 6 evenly spaced planes each possessing 4 evenly spaced satellites. The orbit for each of the satellites is based on the SeaStar satellite, as shown in Fig. 5.2, a satellite tasked with "examining oceanic factors that affect global change" [40]. The SeaStar satellite is a relatively new Earth observing satellite launched by NASA in 1997 and has a primary mission lasting over five years, making it a good candidate to base Earth observing satellite capabilities so that the simulation will be realistic. The SeaStar satellite is in a low Earth orbit, at an inclination of 98.2 degrees [38]. This orbit permits coverage of most of the Earth's surface. The SeaStar satellite has a repeating ground track orbit, which is designed to allow a specific target on the ground to be revisited every 16 days without the need of performing orbital maneuvers. A summary of pertinent SeaStar orbital and repeat ground track parameters is listed below in Table 5.3.

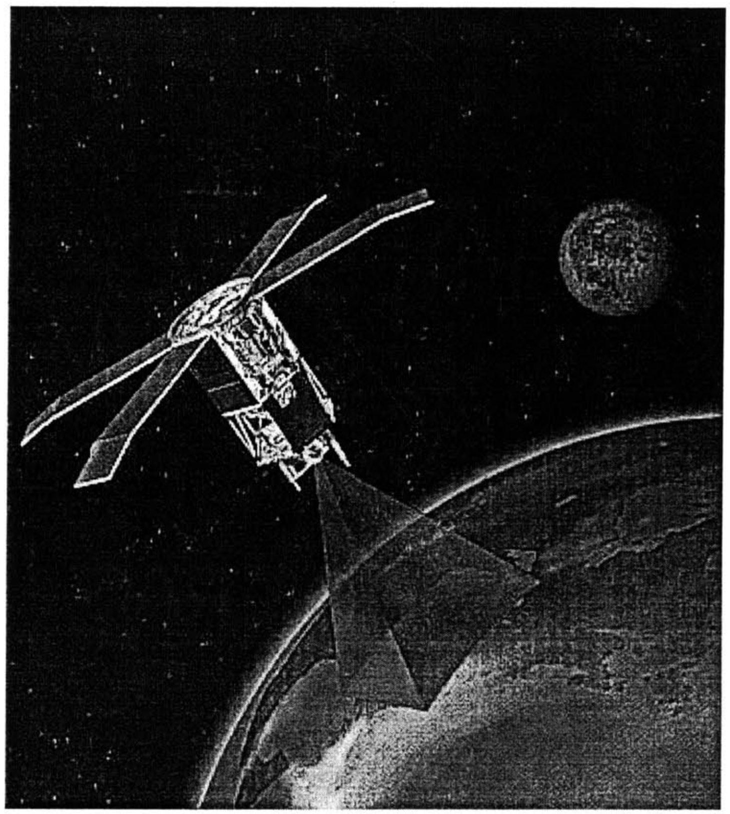


Fig 5.3. SeaStar satellite and sensor footprint.

Table 5.3. List of pertinent SeaStar orbital, repeat ground track and sensor parameters [38].

Orbital Elements and Sensor Parameters	Values
Average Altitude	699 km
Inclination	98.2 degrees
Right Ascension of the Ascending Node	219.2 degrees
Eccentricity	8.25e-5
Argument of Perigee	42.7 degrees
Mean Anomaly	317.4 degrees
Mean Motion	14.58 revs/day
Sensor Field of View	116.6 degrees

The satellites used all possess a single fixed, nadir-pointing sensor. The sensor field of view is based on the field of view that is available with SeaWiFS [35], the sensor onboard the SeaStar satellite. To add heterogeneity in the satellite group, two different sensor types were created, these being a visual and an infrared sensor. While the field of view on both sensors was set to the same value, it was assumed that the actual use of each sensor type would depend on the type of information desired from each target or the circumstances surrounding the target. For example, if visual data is desired from a target, then a satellite with a visual sensor must be scheduled to observe the target. All satellites with infrared sensors would be excluded from being included into the plan. This means that if a satellite with a visual sensor experiences a failure and another satellite must be tasked to replace it, then only satellites with visual sensors could be used for this task. The same situation would occur with satellites possessing an infrared sensor. However, in certain situations a satellite with a different type of sensor as the satellite that has failed could be utilized. One example that was looked at was in the event of cloud coverage. If a satellite with a visual sensor cannot observe a target because of clouds interfering in the data collection process, a satellite with an infrared sensor onboard could be tasked to observe the target, as infrared sensors are not affected by the cloud coverage. This makes use of the assumption that data gathered with the infrared sensor, though not what was originally desired, is better than no data at all.

Each satellite is allocated a certain amount of fuel and is permitted to expend that fuel to perform orbital maneuvers. The types of orbital maneuvers that are allowed are restricted to in plane burns. This restriction eliminates costly out of plane burns and attempts to promote a more efficient use of the satellite's scare fuel resources. The size of the in plane orbital burn is also restricted so that the orbit that is achieved after performing the burn results in a pre-specified repeat ground track. By restricting the satellite to only be placed in orbits that produce a repeat ground track, the satellites can over fly the same target numerous times in the future just by coasting. This is also an attempt to achieve increased observation time by expending only a limited amount of fuel. The repeat ground track orbits that are permitted are listed below in Table 5.4. The nominal repeat ground track is highlighted. This ground track is the actual one being flown by the SeaStar spacecraft and is also the initial ground track used in the simulation.

It is assumed, for simplicity in calculations, that all orbital burns are purely impulsive and occur only at the equator.

Table 5.4. Listing of permissible repeat ground tracks with nominal values highlighted.

Number of Days until Satellite Repeats Ground Track	Number of Revolutions before Satellite Repeats Ground Track
6	91
7	106
1	15
8	119
7	104
6	89
5	74
4	59
7	103
3	44
8	117
5	73
7	102
16	**233**
2	29
7	101
5	72
8	115
3	43
7	100
4	57
5	71
6	85
7	99
8	113
1	14

Targets

Targets that were used in the simulation were chosen to approximate phenomenon that are of interest on Earth today. While targets could have spanned any location or number on the globe, for simplicity sake, only two targets were used. The two targets were of similar types, being an Atlantic hurricane and a Pacific typhoon, and only varied in geographic position. While the position of both of these targets will vary over time, all calculations ma7de with the integrated planner assume that the targets are stationary, for simplicity in calculations. Future work however should examine mobile targets. It is also assumed that the same types of information is desired to be collected from each target as they are the same type of weather phenomenon.

Unexpected Events

Three types of unexpected events were identified, these being satellite failures, changing observation requirements of targets or new opportunities. These three unexpected events were described in more detail earlier in the chapter. Only two of these three were examined for this simulation, as new events occurring after the start of the mission were not studied.

For satellite failures, two types of failures were identified: one, failures that could not be recovered from and two, failures that the satellite could repair itself. If a satellite experiences the first type of failure, other satellites in the group possessing the same type of sensor and not already tasked immediately activate the reaction planner. This results in a satellite being tasked to replace the failed satellite, if one is available. If a satellite experiences the second type of failure, the reaction planner is also activated by other satellites with the same sensor. Because these satellites will not immediately re-task themselves as a replacement, the failed satellite has some amount of time to attempt to repair itself. If the repair is successful in a certain amount of time, the other satellites will not attempt to replace it and will "stand down". If the satellite cannot repair itself or cannot repair itself in a given amount of time, one of the other satellites will be re-tasked to replace the failed satellite, in the same manner as when a satellite experiences a complete failure.

In order for the above procedure to operate, communications between satellites must be provided at some regular time interval. This communication allows satellites to share information with the group on their mission status and health status. If a satellite knows that it is experiencing a failure it can inform the group so that other satellites can take over its tasks. If the satellite experiences a failure and is not able to communicate this with the group, there is an additional safety protocol in the behavioral planner that will still allow other satellites to be re-tasked in the event that communications from a satellite is not received after a set amount of time. This allows a satellite experiencing a complete failure to still have its tasks taken over by another group member without having to possess the ability to communicate the failure to the rest of the group.

The other type of unexpected event that is studied in the simulation is the presence of unexpected cloud coverage. This event has some similarities and differences with satellite failures explained above. First, it is assumed that cloud coverage only affects satellites with visual sensors. Infrared sensors are unaffected by cloud coverage at all. Second, it is constrained that only satellites with infrared sensors are allowed to be re-tasked when a satellite with a visual sensor is obstructed by cloud coverage. This allows observations to still be performed on the target. Two of the most important differences between cloud coverage and satellite failures is that cloud coverage can be predicted in a limited sense and that the coverage is temporal. It is assumed that cloud coverage can be predicted based on weather forecasts before the satellite encounters the target and discovers its sensor is obstructed by clouds. It is assumed for this simulation that cloud coverage can be predicted up to 2 hours before the satellite observes the target. This prediction capability allows additional time in re-tasking other satellites so that it may be possible to observe the target near the time originally planned for. Without this prediction capability, there would be no possibility of observing the target during the time planned. Because cloud coverage is temporal, there is no certainty that if a satellite with an infrared sensor is tasked to observe the target the clouds will not soon dissipate and render the re-tasking pointless, wasting precious fuel resources. For this reason, a satellite must have a prediction of losing more than one viewing opportunity before a new satellite is tasked. If a satellite predicts that it will lose a second viewing opportunity in a row, it will request that another satellite be tasked to observe the target. For this simulation, it was assumed that all predictions are deterministic, meaning that all are correct 100 percent of the time. Future work could include predictions that are stochastic in nature.

5.4 Scenarios

The scenarios that were constructed had the dual goal of providing a set of circumstances to test the various aspects of the integrated planner and to allow testing to be conducted that would determine whether the use of the reaction planner was beneficial if an unexpected event occurred. This section provides an overview of the scenarios that were constructed and presents the items of interest that were studied for each scenario.

Scenario Overview

Two major types of scenarios were constructed, the first of which is a scenario that will subject the satellite group to a single failure. The second type of scenario subjects the satellite group to a set of multiple failures during the course of the mission. The first scenario type allows the study of the benefit derived from the reaction planner's use while the second type of scenario allows the reaction planner's capabilities to be more fully tested. The multi-failure scenarios allow the reaction planner's ability to respond to multiple failures happening either in series or in combination.

Several assumptions were made for all the scenarios. First all satellites in the group were assumed to be in the same orbit, but at different phases. The sensors that each satellite had on board all had the same field of view, even though the sensor types were different. The number of satellites having each sensor type on board was evenly split among the group, with half possessing the visual sensors and the other half possessing the infrared sensors. The placement of the sensors in the Walker constellation was also evenly distributed, with alternating satellites possessing the different sensors. Although the integrated planner has the capability to create plans for multiple targets, one target was concentrated on for most scenarios. This was because the observation plans that are created by the integrated planner are almost completely decoupled, with the only link between the two the non-availability of satellites that are being used to view one target to be used in a plan for the second target. Concentrating on only one target allows few satellites to be used in the group to generate the same plan types, allowing the simulation to run faster. Another assumption is that the original plan that is generated by the EPOS 1.0 optimal planner will remain in effect, as much as is possible, after the presence of a failure. It is assumed that there will be additional satellites that can be re-tasked to try and repair the original plan that was damaged through the loss of the failed satellite. This assumption stems from an assumption made from the EPOS 1.0 system, which is that there is no automated attempt to create an optimal group plan. All optimal plans that are generated are done at the satellite level. Any attempt at creating an optimal group plan, or combining all the individual plans, must be accomplished manually by the user. Also stemming from an assumption made in EPOS 1.0 is that there is no provision to match the actual coverage dates of the original plan with any plan that is generated in the integrated planner. Currently there is no automatic procedure to optimize the observation plan to include specific viewing dates. The assumption was made when re-tasking new satellites that the satellite that would provide the greatest amount of observation coverage of the target was the best satellite to re-task, if available. No attempt was made to limit the amount of fuel that the re-tasked satellite would use beyond that which is limited by the EPOS 1.0 optimizer.

Parameters of Interest

Several parameters of interest were identified that were studied when using the integrated planner. The parameters identified can be divided into three major categories; benefits, costs and operational conditions. A summary and brief explanation of all the specific parameters that were studied is provided below in Table 5.5.

Table 5.5. Summary and explanation of parameters of interest.

Parameters of Interest	Explanation of Parameters
Benefits	Benefits are measured as total observation time that the satellite can achieve of the target
Total group benefits	The sum of all satellite's total observation time of the target that have been tasked to view the target
Lost group benefits	The sum of all satellite's observation time that have been tasked to see the target and experience a failure, summed from the start of the failure to the end of the mission
Additional group benefits	The sum of all satellite's observation time that have been re-tasked to observe the target
Costs	Costs are measured as either total fuel used (measured in delta V) or as number of satellites used
Total fuel expended	The sum of all satellite's total fuel used in orbital maneuvers
Additional fuel expended	The sum of all satellite's fuel used in orbital maneuvers that have been re-tasked to view the target
Total satellites tasked	The number of satellites that have been originally tasked to view the target
Additional satellites tasked	The number of satellites that have been re-tasked to view the target
Benefit to Cost Ratio	The ratio of benefit to cost (measured in delta V or satellites) that is achieved by the group
Operational Conditions	Parameters that affect the benefits and costs when using the integrated planner
Number of satellites in group	The total number of satellites available in the group to be used for any plan or re-tasking
Total satellites available	See above
Ratio of satellites used in original plan to satellites available	The ratio of satellites that are originally used to the ratio of satellites available
Ratio of satellite failures to number of satellites in original plan	The ratio of satellite that fail to the number of satellites that are in the original plan
Mission time	The length of time that a particular target is of interest
Length of mission	See above
Ratio of time to failure to total mission time	The ratio of time remaining between when a failure occurs and how much time is left to recover from the failure by the reaction planner

5.5 Integrated Planner Example

The following section contains an example that serves to illustrate how the integrated planner runs and what types of results can be expected as well as what type of information is required to operate the integrated planner. Included are several captions taken from the integrated planner and a corresponding explanation. As shown above in Fig. 5.1, the integrated planner consists of two major sub-systems, EPOS 1.0 and the reaction planner. EPOS 1.0 operates entirely before the mission begins, while the reaction planner only operates if there is an unexpected event during the course of the mission.

User Inputs

To begin the planning process several pieces of information are required by the integrated planner. This information can be subdivided into two major categories; one that is mission specific and the other that is parameters that will affect the operation of the satellite group. Some examples of the first type of data are number of satellites available, the length of the mission and the target of interest. Examples of the second type of data include the cap on fuel that each satellite is allowed to expend, the rate of communication broadcasts between each satellite and the amount of time that cloud coverage can be reliably predicted. This section concentrates on covering the first type of user input, which is more mission specific, while the second type of input is believed to be more hardware specific.

When starting the integrated planner, the user is presented with a graphical user interface, or GUI, in which they can input mission specific data. The initial GUI that is presented to the user is shown below in Fig 5.3.

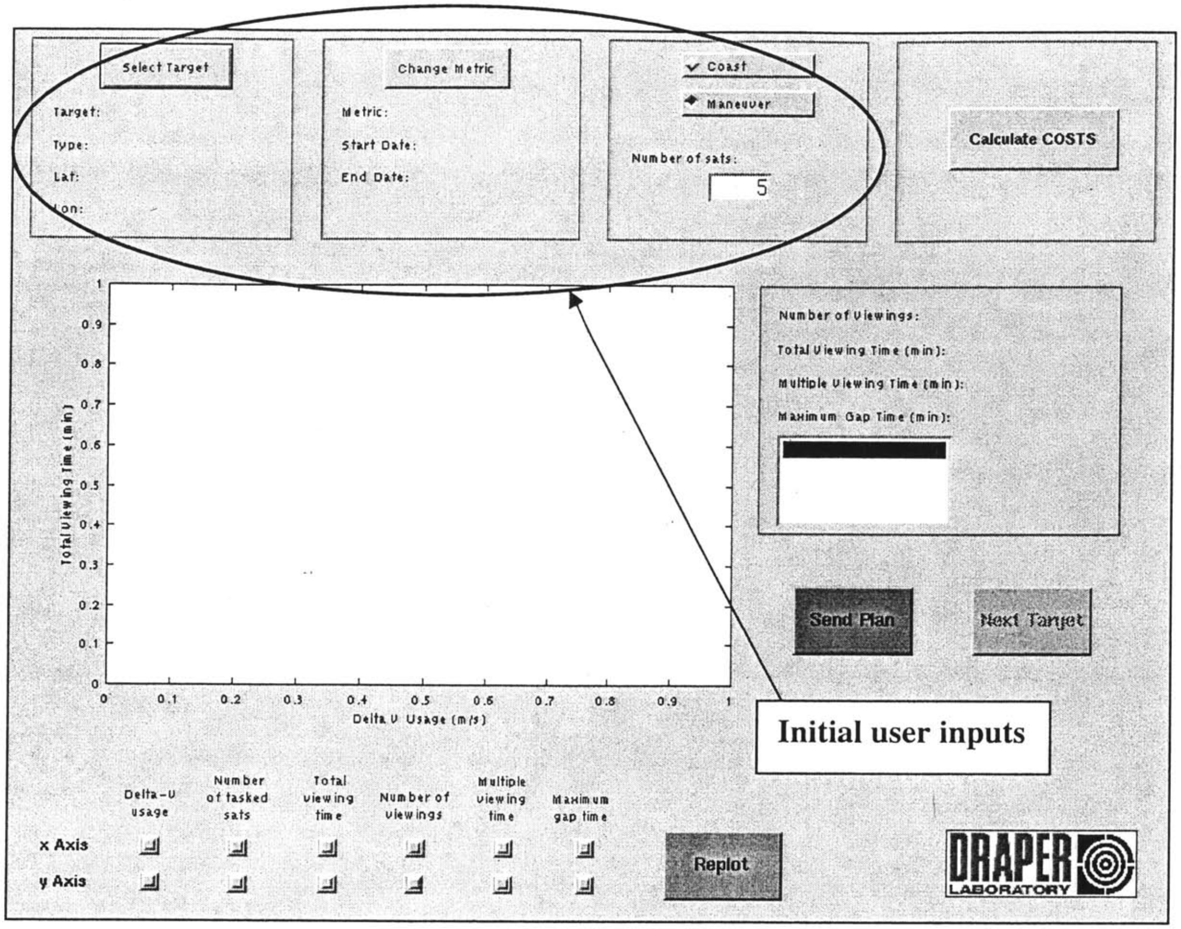


Fig 5.4. Initial GUI Interface for user inputs of integrated planner.

The initial inputs that are required from the user are information concerning the target that is desired to be observed, information on the mission length, the metric that will be used to allocate satellite resources, the number of satellites available and either allowing the satellites to perform orbital maneuvers or restricting satellites to only coasting. The information that is required for each of these inputs are illustrated below in the following figures.

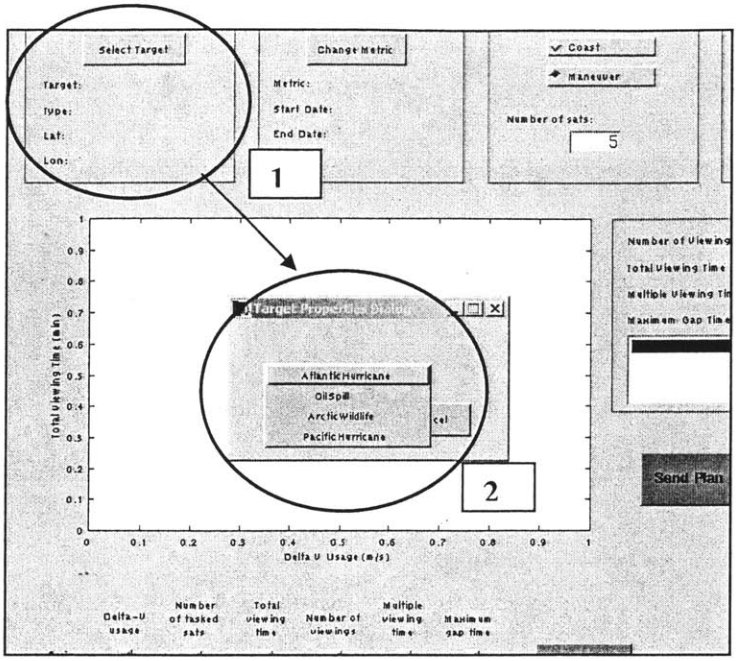


Fig. 5.5. User target selection.

Fig. 5.4 illustrates the user's ability to choose a target from a list of available targets, as shown in the second bubble. The available targets for this sample run include hurricanes in the Atlantic and Pacific oceans, an oil spill off the coast of Alaska and the Artic Wildlife Preserve in northern Alaska. Additional targets would require information to be added to EPOS 1.0. After the target has been selected, a summary of the information is presented in the first bubble, as shown in Fig. 5.6. The target which has been selected is the Atlantic Hurricane, which is a hurricane off the eastern coast of Florida.

The next set of data that the user is prompted to supply the integrated planner is the desired metric and the start and end dates over which the mission will begin and end. Several metrics are available for selection, as shown in the fourth bubble, including concentrating on achieving observations as soon as possible, maximizing the total number of observations or maximizing the total time that the target is observed. Additional metrics would require information to be added to EPOS 1.0. After the metric and the mission time horizon has been selected, as summary of the information is presented, as shown in the third bubble in Fig. 5.7. The maximize total observation time has been chosen for the metric over a mission time horizon lasting two days starting on July 4, 2001.

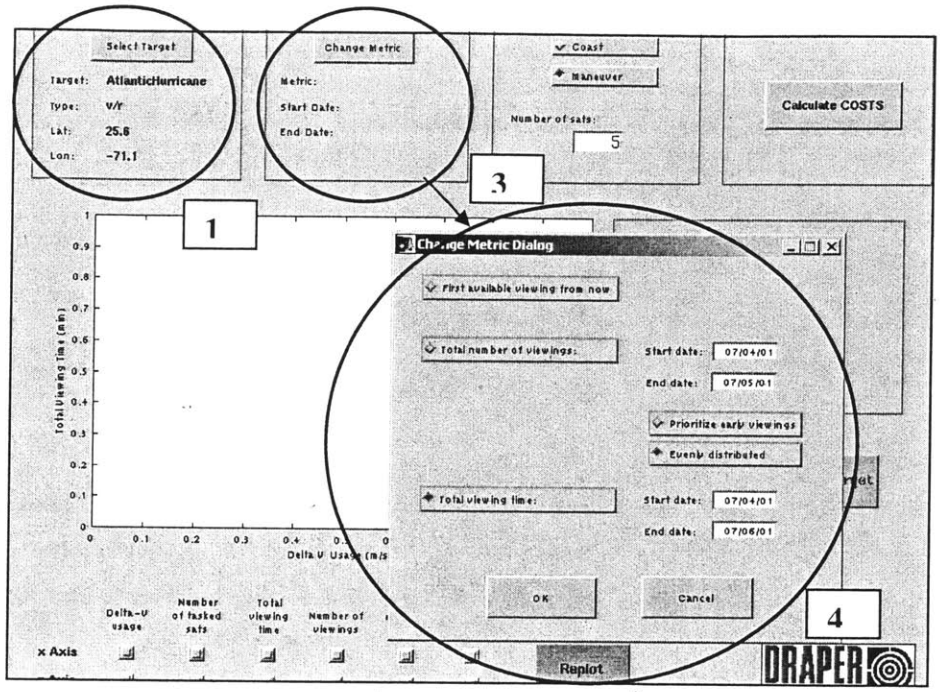


Fig 5.6. User metric and mission time horizon selection.

The last set of initial information that the user is prompted for is the decision on whether to allow targets to perform orbital maneuvers or coast and to determine how many satellites are available for the mission. For this example the satellites have been allowed to perform orbital maneuvers and eight satellites have been allocated to observe the target, as depicted in the fifth bubble in Fig. 5.7. A depiction of the GUI just before the information is set to the optimal planner is shown below in Fig 5.7. Pressing the "Calculate Costs" button, as shown in the sixth bubble in Fig. 5.7 sends all user selected information to the optimal planner.

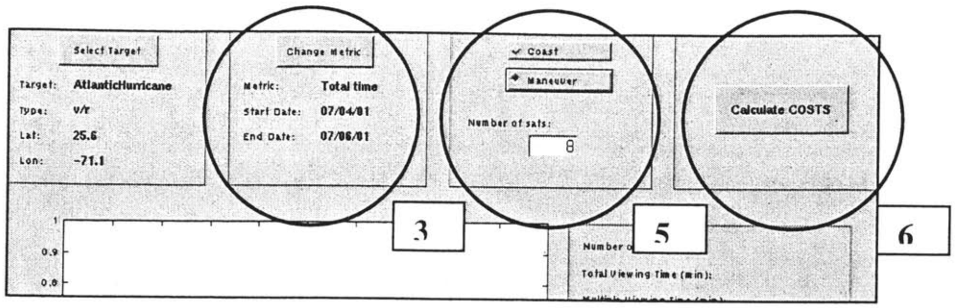


Fig. 5.7. User selection for orbital maneuvers and number of available satellites.

Once the optimal planner in EPOS 1.0 has determined the optimal trade-off between observation time and fuel usage for each satellite, the user is presented with this information in graphical form, as shown below in bubble seven in Fig. 5.8. The choice of which set of satellites to actually task from the total group is a decision that the user must make, augmented by the information supplied. For this example a total of eight satellites are available, but it may not be desired to use all eight when observing the target for a variety of reasons including that the level of coverage that could be attained from using all eight satellites is not necessary, that the cost involved in using all eight is too high, that some satellites are being reserved in case another target appears or to hold some satellites in reserve in the event that one of the satellites fails and must be replaced. The user must determine how many satellites, from one to eight, and which satellites should be used. As this is a combinatorial problem, the number of different possibilities increases exponentially as the total number of satellites in the group increases. If the number of possible satellite set combinations exceeds a given threshold, EPOS 1.0 will automatically eliminate some of the satellite sets that have the poorest performance. The user is still presented with the remaining sets and must make the final decision.

For this example three out of the eight available satellites were chosen to actually observe the selected target. This decision was made by limiting the number of satellites that would be tasked to three of the available eight and allowing a moderate amount of fuel to be expended in observing the target, as represented by the red square highlighting the satellite set choice. After applying these criteria, the set of satellites that provided a high amount of observation time was selected. For this example satellites 1, 3 and 6 were selected.

Additional criteria could be selected to make the decision, if needed. Some potentially important considerations that the user may find important are anticipated and provided to the user when making a decision. These are shown below in Fig. 5.8 in bubble eight and include such parameters as maximum gap time between satellite observations, the total number of satellite over flights that the group achieves and the number of viewings that are performed simultaneously by more than one satellite. Selecting any of these options re-plots the information presented in the GUI. One such possibility is shown below in Fig. 5.9, in which the total number of viewings as a function of fuel used is presented.

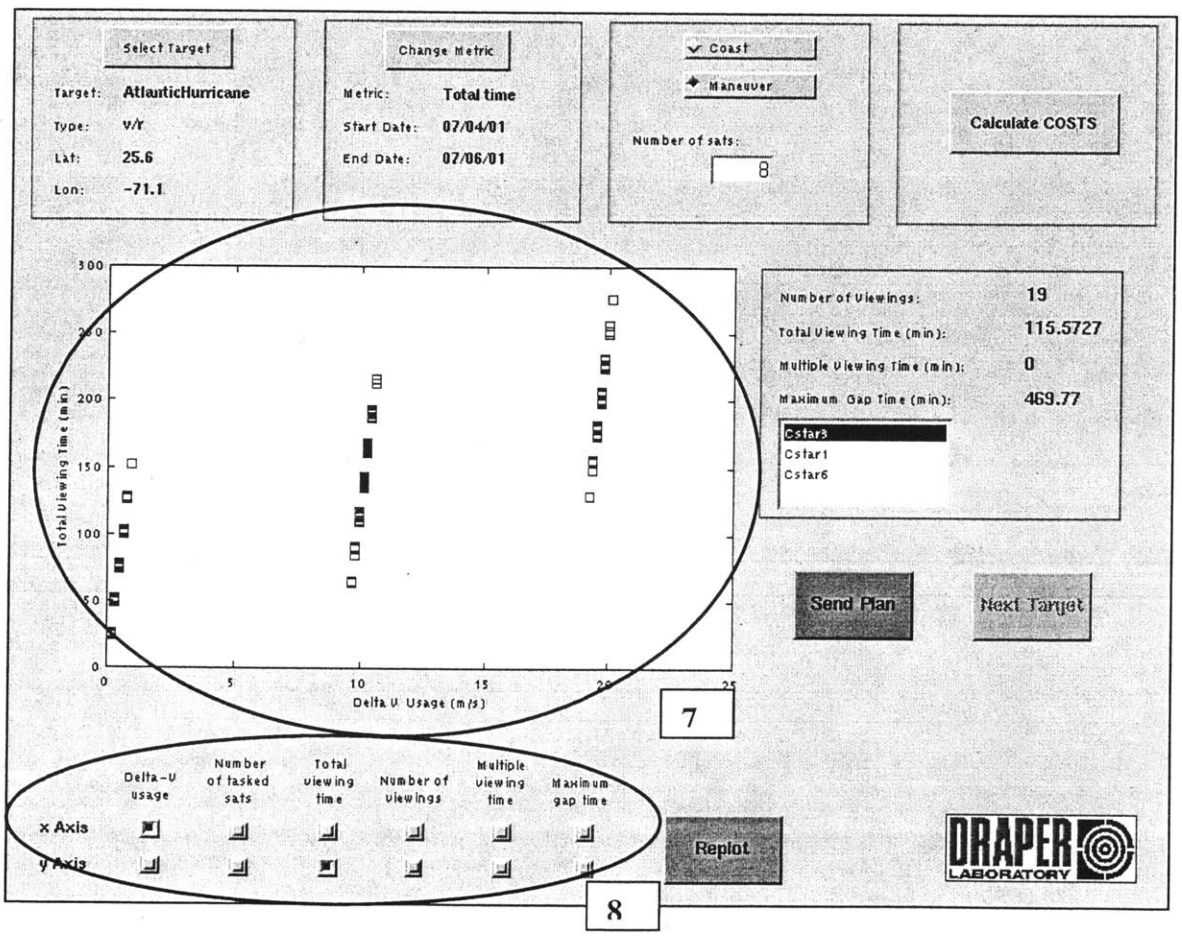


Fig. 5.8. Possible satellite sets that can be tasked.

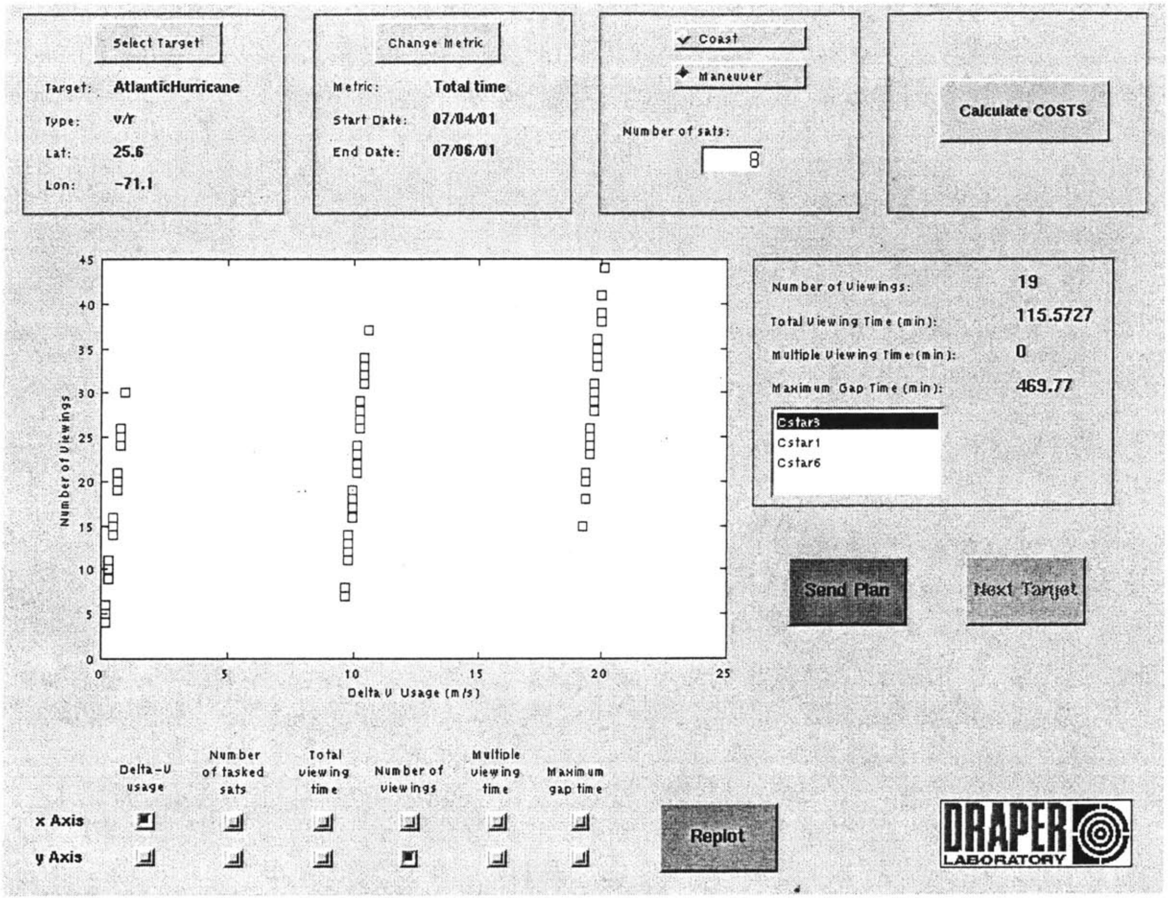


Fig. 5.9. Possible satellite sets that can be tasked, presented with alternative information.

Once the user has selected the preferred satellite set, pressing the "Send Plan" button informs EPOS 1.0 that a maneuver and observation plan can be constructed for the satellite set. This information would then be uplinked to the satellites in the set and once the mission time is reached the satellites would be free to begin performing the orbital maneuvers and observing the desired target.

For this example, three unexpected events are assumed to occur to the three satellites that have been tasked. Satellite 1 has a visual sensor onboard and at some time in the mission runs into cloud cover that is predicted to last long enough to obstruct multiple sequential viewings. Satellite 3 also has a visual sensor onboard and experiences a total failure.

Satellite 6 has an infrared sensor and experiences a failure, but is able to repair it only to experience the same failure again at a later date that it cannot recover from. Once the mission begins and a failure is detected the reaction planner is engaged. The reaction planner selects a satellite, if available, to replace the failed satellite according to the procedure overviewed earlier in this chapter and the previous chapter, concerning the model based predictor and behavioral planner. Below is the output from the reaction planner over the course of the mission, showing each satellite's motivation over time.

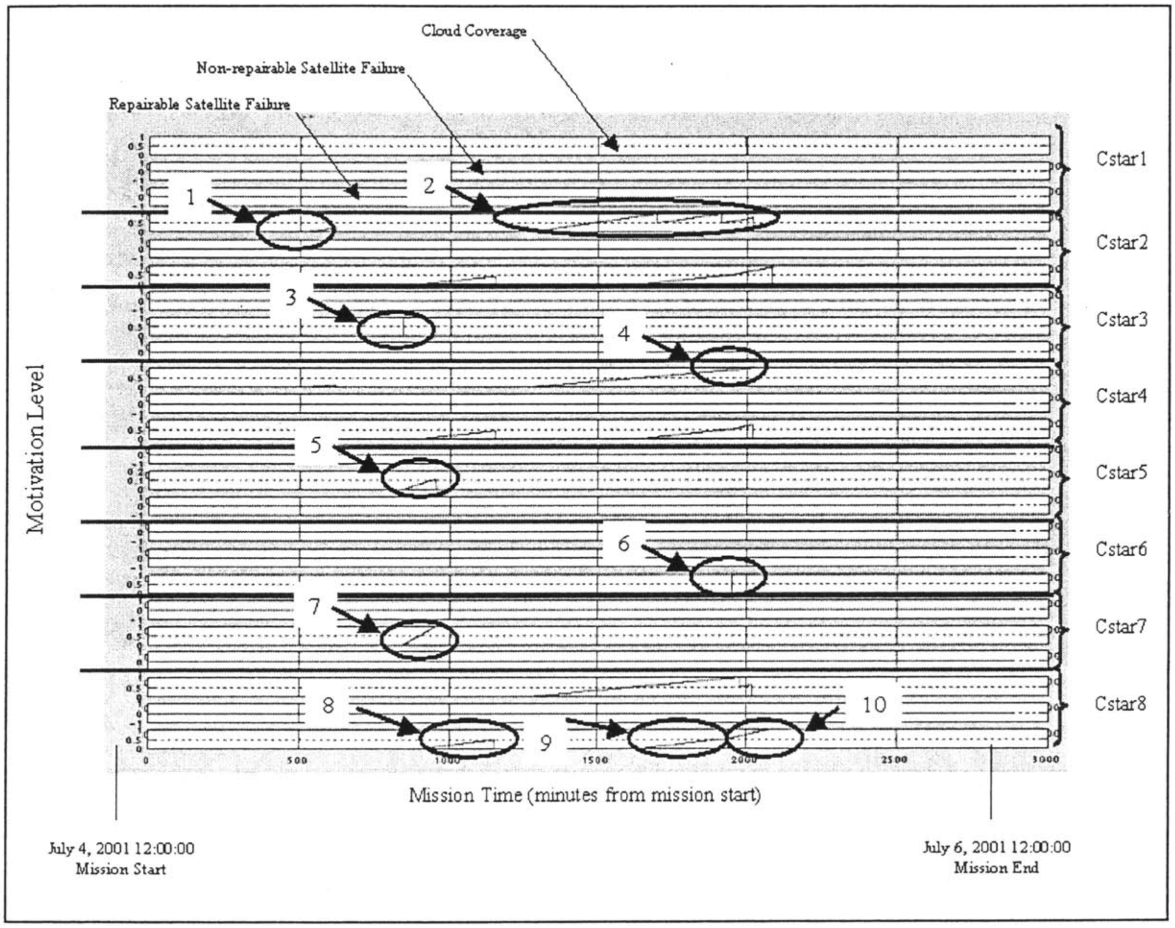


Fig. 5.10. Output from the reaction planner, motivation for each satellite as a function of time.

Figure 5.9 displays the motivation history for a mission with eight satellites experiencing three types of unexpected events. The odd satellites have visual sensors onboard while the even satellites are equipped with infrared sensors. To limit the amount of data that is presented, only the motivation history for the eight satellites as they relate to the unexpected events are presented. The reaction that the group has to these events is explained below and referenced back to the above figure.

At the start of the mission, Satellites 1, 3 and 6 have been assigned to observe a selected target. The rest of the satellites in the group of eight have not been assigned any tasks. At time zero, the time when the mission begins, the motivation level for Satellites 1, 3 and 6 is at its threshold value of one, meaning that these satellites are working on an assigned task. The motivation levels for the rest of the satellites are at zero, because they have not been assigned a task and there have been no unexpected events.

The first unexpected event is cloud coverage over a target that Satellite 1 is tasked to observe. As Satellite 1 possesses a visual sensor, cloud coverage will prevent it from completing its mission. However, to keep satellites from being re-tasked every time there is cloud coverage, satellites will be re-tasked only when there is sustained cloud coverage, where sustained cloud coverage is defined as cloud coverage over tow or more consecutive viewing opportunities. The first time cloud coverage is observed, available infrared satellites begin to increase their impatience, as shown in bubble 1 for Satellite 2, with a similar rise for Satellites 4 and 8. As the next viewing opportunity affords clear skies, enabling Satellite 1 to again view the target, the motivation levels in Satellites 2, 4 and 8 drop to zero.

Cloudy skies again threaten Satellite 1 around time 1250. This time two observation opportunities in a row are obscured and an infrared Satellite is re-tasked to replace Satellite 1. The satellite that replaces Satellite 1 is Satellite 4, as shown in bubble 4.

The second unexpected event that occurs is a failure in Satellite 3, as shown in bubble 3. This satellite does not posses any self-repair capabilities and thus its motivation to perform a task drops to zero because of the failure. As Satellite 3 is a visual satellite, another visual satellite is desired. Only Satellites 5 and 7 are available, because Satellite 1 has been tasked at the outset of the mission and cannot be considered. Of Satellites 5 and 7, Satellite 7 is expected to be able to complete a greater number of observations over the course of the mission so its impatience is given a higher weighting. Because of this weighting it reaches a motivation threshold before Satellite 5, as shown in bubble 7, and thus begins the tasks that Satellite 3 was attempting to accomplish before its failure. As a Satellite has been tasked, Satellite 5's motivation level returns to zero, as shown in bubble 5.

The third unexpected event that occurs is a failure in Satellite 6 around time 900. While this does not appear directly on the motivational history for Satellite 6, the failure can be inferred from the motivation levels of the rest of the group. As Satellite 6 has an infrared sensor, a satellite with an infrared sensor is desired to replace it. The other satellites with infrared sensors in the group are Satellite 2, 4 and 8. After the failure in Satellite 6 is reported to the rest of the group, the impatience level raises in Satellite 2, 4 and 8, as called out in bubble 8 and shown as the similar rise in motivations for Satellites 2 and 4. The rise in motivation level is halted however around time 1200, as Satellite 6 has the ability to perform self-repairs. Around time 1200 it completes self-repairs successfully and transmits this information to the rest of the group, causing the motivation level to drop to zero again for Satellite 2, 4 and 8.

Around time 1700 Satellite 6 again experiences a failure. As it attempts to repair itself, the impatience level in other infrared satellites slowly rises, as shown in bubble 9. However, this time Satellite 6 cannot repair itself in time and the impatience level of the other satellites increases at a faster rate, as shown in bubble 9. As Satellite 6 still cannot complete repairs, Satellite 8 is re-tasked to replace Satellite 6, as shown in bubble 10 and 9.

One interesting anomaly that is apparent in this simulation is shown in bubbles 2 and 4. These bubbles highlight the increasing motivation levels of Satellites 2 and 4. It is apparent from the figure that Satellite 2 is better suited to be re-tasked. This is apparent because its impatience rises at a faster rate, as illustrated by the fact that the motivation level reaches the threshold several times while Satellite 4 has a lower motivation level.

Satellite 2 is not allowed to be re-tasked at these points where it reaches the threshold because it is not yet certain that the cloud coverage that is affecting Satellite 1 and driving Satellites 2 and 4 will still be a problem when Satellite 1 again passes over the target. Until it is certain that Satellite 2 or 4 should be re-tasked, its motivation will be decreased by half if it reaches a threshold. The motivation is allowed to continue growing and for Satellite 2 reaches the threshold again and is again reduced by half. The anomaly occurs when it is determined with certainty that a new satellite is needed and the satellites are allowed to come up to the motivation threshold without being reduced. By coincidence, the timing and motivation values for Satellites 2 and 4 allow Satellite 4 to reach the threshold value before Satellite 2, even though Satellite 2 is better suited for the task and has a faster growing impatience. This results in a satellite being tasked that is not the best satellite available. This is a flaw in the current formulation, likely attributed to selection of parameter values and should be looked into further in the future.

The output that is generated from the integrated planner can be viewed in animated format in the commercial off the shelf software package Satellite Toolkit, or STK. This allows the user to simulate what would actually happen if the integrated planner were implemented with satellites that were in orbit. A sample graphic of the animation provided from STK is presented below in Fig. 5.10. The time displayed in this figure is before any of the satellites have experienced an unexpected event. Satellite 6 is shown to be just finishing its over flight of the Atlantic Hurricane target.

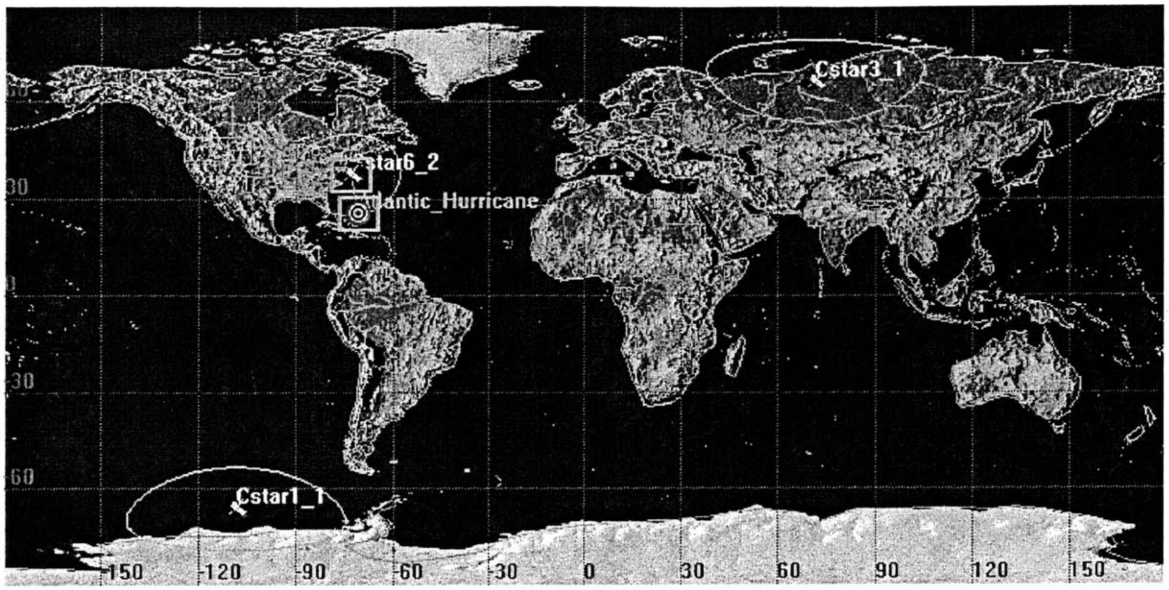


Fig. 5.11. Sample animation frame showing satellite overflying the target in STK.

Chapter 6

Results

The results chapter is divided into an overview of the results, a detailed discussion of the figures, produced from data generated while using the integrated planner and a summary of the results. The detailed discussion is further divided into results pertaining to the need for orbital maneuvers, the benefit of using the reaction planner, the cost of using the reaction planner, a comparison of the benefits to the costs of using the reaction planner and expected performance levels from the reaction planner.

6.1 Overview of Results

The detailed results that will be presented in the following section offer several interesting observations on the use of the integrated planner that could be applicable to understanding the performance characteristics of dynamic satellite constellations. This section provides an overview of these results.

Guarantee of Reaction Planner

The reaction planner that was developed does not provide any guarantee that an optimal agent for the task will be chosen or that an agent will be able to respond to a needed task. During the use of this algorithm however, a satellite would always be tasked, as long as a satellite was available to be tasked, and there was enough time left in the mission to perform orbital maneuvers. However, because the information supplied to the reaction planner was based on outdated information, a satellite was occasionally tasked that could not perform the observation. This occurred only near the end of the mission when the reaction planner tasked a satellite to observe a target because previous data showed that it could achieve an observation, only to find that when the satellite's updated orbital information was examined that the amount of fuel necessary to perform the maneuver would have been too great. The result was that a satellite was tasked but was not allowed to perform the orbital maneuver and achieve the observation because of the large fuel cost.

The optimal selection of an agent was also not guaranteed by the reaction planner. An attempt was made to influence the tasking of satellites that would achieve the maximum amount of observation time possible for the remainder of the mission. Two problems were in evidence that sometimes prohibited the best agent from being selected. First, the decision was made on outdated information. If the information was updated throughout the mission, this could be eliminated as a reason for the optimal satellite not being selected. The other reason was implementation of the algorithms. It was noticed on multiple occasions that due to varying circumstances, the optimal agent was not selected. This effect is explained in Chapter 5 in greater detail. It is not known whether this is a fault of the ALLIANCE algorithms or the implementation of the algorithms by the author. Either way, it is anticipated that this effect could be removed with relatively little effort. It was not worked with due to the low frequency of occurrence that was observed.

Constellation Design

As was discussed in the detailed results section, the position that satellites have with respect to one another greatly influences the ability for the constellation to effectively respond to unexpected events. It was seen that satellites that were synchronized with one another had the best percent of observation time regained after an unexpected event. This came, however, at the cost of using the largest amount of fuel, on average, of any satellite in most situations. Also, these satellites were the slowest to respond after an unexpected event. At the other end of the spectrum, satellites that are in or near the same plane as the satellite that is subjected to the unexpected event have the shortest response time near the original observation time. The fuel used by these satellites and the percent observation time achieved are both between that used and achieved by the synchronized and non-synchronized satellites. At first glance these satellites would appear to be a good compromise in terms of performance and cost. However, being in or near the same plane as the original satellite may limit the utility of these satellites when they are not needed to respond to an unexpected event. While they can increase coverage of a particular target, it is hard to cover multiple targets simultaneously with satellites that are all in or near the same plane.

These results suggest that dynamic constellations should be built specifically for each mission and should be designed according to what is expected, in terms of observations and need to respond to events that cause a loss of planned observation time. It is thought that if the resources are available, the probability is large that unexpected events will occur and the need to regain lost planned observation time large, that a dynamic constellation consisting of satellites from all three types is desirable. If the resources to do this are not available, it is thought that a set of non-synchronized satellites is the best compromise for most missions. This is because multiple targets can be viewed simultaneously, the same target can be viewed with spaced frequency, and the ability to recover most observations after the occurrence of an unexpected event is possible.

Benefit to Cost Ratio

All the results in this thesis have been generated using a metric that sought to maximize observation time over the course of the mission. This resulted in several orbital maneuvers being performed that expended a large amount of fuel and provided only a very limited amount of observation in return. While maximizing observation time is a viable metric to use to plan observation and maneuver schedules, other metrics are available. One such metric is to chose satellites based on their benefit to cost ratio instead of the total observation time achieved. Some preliminary trails of using benefit to cost ratios as opposed to maximum observation time as the metric of choice gave large increases in fuel efficiency while sacrificing only a small fraction of the total observation time achieved. Benefit to cost ration increased from around 0.5 min/m/s using the current metric to around 25 min/m/s when using the ratio as the metric. Average observation times achieved decreased only by a few minutes with the new metric. This extreme increase in fuel efficiency points to further thought needing to be given about choice of metrics.

6.2 Detailed Results

A set of detailed results are shown in the following figures, depicting the need for orbital maneuvers, the benefit of using the reaction planner, the cost of using the reaction planner, a comparison of the benefits to the costs of using the reaction planner and expected performance levels from the reaction planner.

Need for Orbital Maneuvers

It has been assumed throughout this thesis that orbital maneuvers are a viable means of increasing the amount of time that a satellite will be able to observe a specific target. This section presents results that show the benefit of using just the EPOS 1.0 optimal planner to plan orbital maneuvers over that which is achievable by restricting the satellites to a static, coasting orbit. The results are of the form of showing viewing time achieved using coasting and maneuvering satellites over a range of different mission lengths.

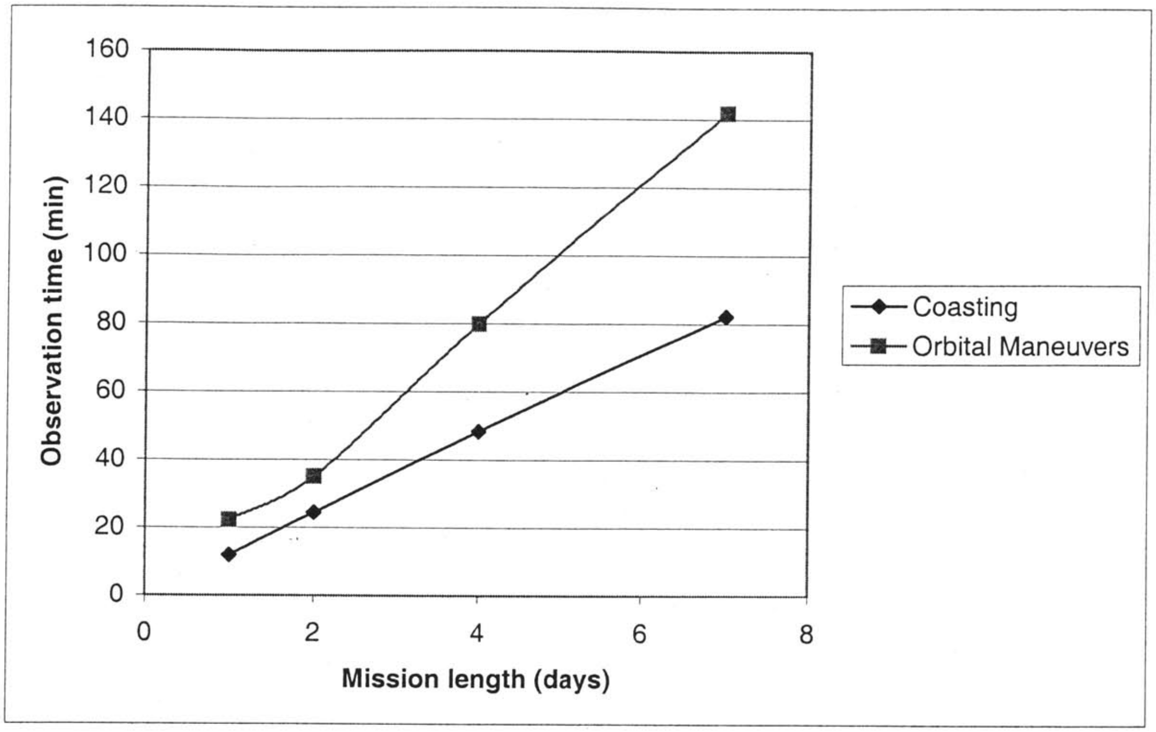


Fig. 6.1. Observation time achieved with coasting and maneuvering satellites.

Fig. 6.1 Discussion

This figure displays the average time that a satellite can observe a specified target. The observation time that is achievable when the satellite is restricted to its initial orbits is displayed, along with the time that is achievable when the same satellite is allowed to perform orbital maneuvers. This amount of observation time is varied as a function of the total length of the mission, with missions ranging from 1 to 7 days examined. For this figure, the average time is calculated by averaging the times achievable from a group of eight satellites.

Fig. 6.1 Observations and Interpretation of Results

- The observation time that is achievable when using coasting satellites and maneuvering satellites increases as the mission length increases. This is an obvious result, as the amount of time that the satellite can physically travel over the target increases over a longer mission time.

- The rate of increase of observation time for the coasting satellites is linear with respect to mission time. This is to be expected as all the satellites are in a repeat ground track orbit which dictates how many times a satellite will pass over a target in a given number of days. If the number of days were to double, the observation time would also double. While the observation may not be strictly linear for an individual satellite, due to the discontinuous nature of observation times, the average of several satellites does produce a very linear increase with respect to mission time.
- The rate of increase of observation time for the maneuvering satellites is almost linear with respect to mission time. The linearity of this trend is due to the fact that satellites were allotted a specified amount of fuel for the entire mission. Satellites did not have to burn all of this fuel, but were allowed to burn as much as needed to maximize the observation time that was possible over the course of the entire mission. This allowed the satellites to increase the amount of observations at the same rate for each increase in mission length. Therefore, the slope of the increase that is seen in the figure is due to the maximum amount of fuel that the satellites were allowed to burn when performing orbital maneuvers. If the amount of fuel were to be increased, the number of maneuvers that could be performed could also increase causing an increase in the total observation time that is achievable. This would show up in the figure as a steeper slope to the maneuvering satellites observation time trend line. Similarly, if the amount of fuel available decreased, the slope would decrease. There are limits in both directions to what slope this line can take. If the amount of fuel were reduced to zero, the satellites would become coasting satellites and the slope would be the same as that shown for the coasting satellites in the figure above. In the other extreme, because the satellites have been restricted to in plane phasing maneuvers, the maximum amount of times that a satellite could observe a target is based on both the number of times that the target passes near the satellite's orbit plane (passes through exactly twice a day) and how large a foot print the sensor on the satellite produces. This restricts the total amount of times that a satellite can view a target in a given day and any additional use of fuel will not reduce any additional observation time of the target.

- There is a small interruption of the linearity of the maneuvering satellites observation trend line. This occurs at the one-day mission length. Because the mission length is so short, it is very "hit or miss" for a satellite to naturally view a target without performing orbital maneuvers. With the amount of fuel allocated to the satellites, each satellite can definitely view the target in a mission of one day's length. This has a slight increase on the benefits of using maneuvering satellites over coasting satellites for short mission times – without maneuvers it is probable that no observation may be possible.

Benefit of Using Reaction Planner

This section presents results that show the benefit of using the reaction planner in conjunction with the EPOS 1.0 optimal planner. Benefits are presented in terms of additional observation time that is achieved by using the reaction planner over that which would be possible if an unexpected event occurred during the course of a mission planned using only the EPOS 1.0 optimal planner. The time during the mission when the use of the reaction planner is a viable option is examined as is the benefit achievable as a function of the amount of time that elapses before a satellite is re-tasked and can view the target after an unexpected event occurs is also presented.

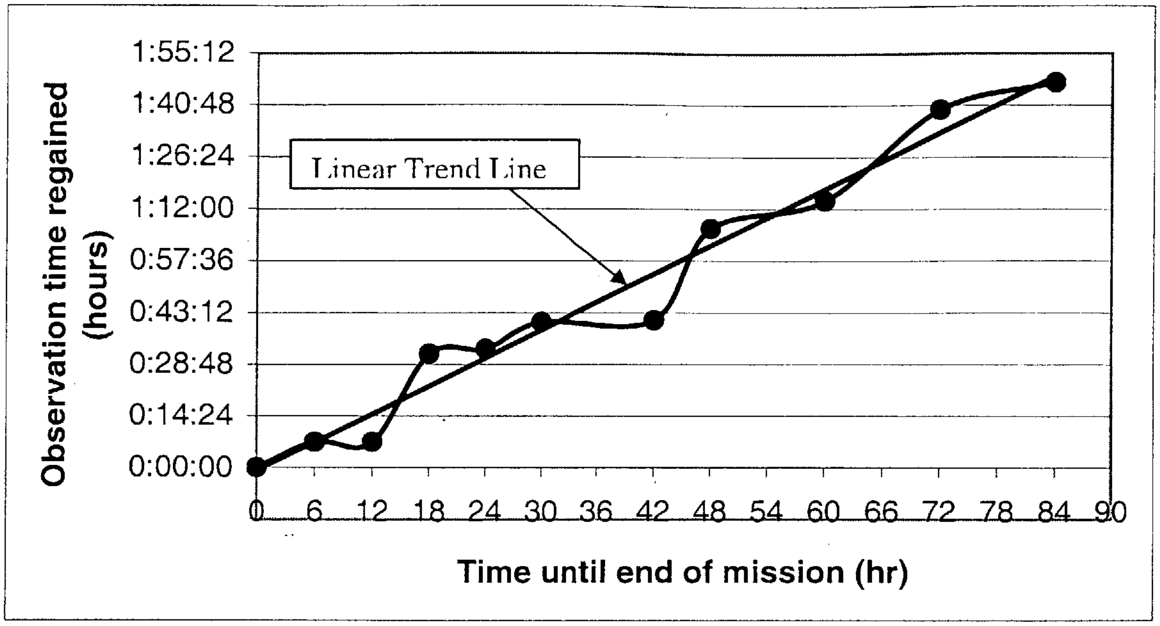


Fig. 6.2. Observation time achieved by re-tasked satellite as function of time from tasking until end of mission.

Fig. 6.2 Discussion

This figure displays the observation time that is regained by a satellite that has been re-tasked to view a target after the occurrence of an unexpected event. The observation time that is achieved is presented as a function of the time that is left in the mission when the unexpected event occurs. A linear trend line is fit to the data produced. Observation time was observed for unexpected events occurring between 0 and 84 hours from the end of the mission.

Fig 6.2 Observations and Interpretation of Results

- The observation time that is achievable with the re-tasked satellite increases as the length of time that is available in the mission after the unexpected event occurs increases. This is expected, as there is more time for the satellite to view the target when there is more time remaining in the mission.
- While the trend of increasing observation time with increasing mission time remaining is close to be linear, it is not exactly linear. The difference between this figure and Fig. 6.1, which was much more linear, is the time interval that data was

taken. In this figure, observation time was measured at time steps down to every six hours, while for Fig. 6.1, the time step was measured on the order of 1 day. The shorter time step produces a much more discontinuous trend, as observation time is only increased when the satellite passes over the target, which is a discontinuous function. This is because the time that a satellite will pass over a target will not be spaced equally throughout the day. For example, a satellite may pass over a target two times in a row at the beginning of the day, and then not pass over the target again for the remainder of the day, because of the constantly changing relative positioning of the satellite and the earth fixed target. While this will average out for the larger time steps, it appears as a series of step increases in observation time when the time step is finer. Practically, this may mean that even if the amount of time increases, the amount of observation time that is possible may or may not increase. For example, as seen from the figure, if an unexpected event occurred 24 hours from the end of the mission, no additional view time would be possible than if the unexpected event had occurred with just 18 hours of time left in the mission. Practically speaking, this has an implication on estimating precisely how much observation time will be achievable with any given satellite.

- It was seen that as the mission neared the end, the observation time that was achievable with a re-tasked satellite approached zero. At the extreme, if there is no time left in the mission, there is no observation time that can be achieved by a re-tasked satellite. Because of the discontinuous nature of the observation time achievable as a function of mission time remaining, it was found that when an unexpected event occurred less than six hours until the end of the mission, it was very difficult to achieve any additional observation time with a re-tasked satellite. This is because the amount of fuel that the re-tasked satellite would need to expend would be greater than the amount allocated to it when performing the large orbital maneuvers necessary to change its orbit drastically in a short amount of time. Practically speaking, this limits the effectiveness of using any type of satellite re-tasking near the end of a mission. This means that an unexpected event near the end of a mission may not be able to be recovered from. This effect, and methods to eliminate it, will be discussed later in this chapter.

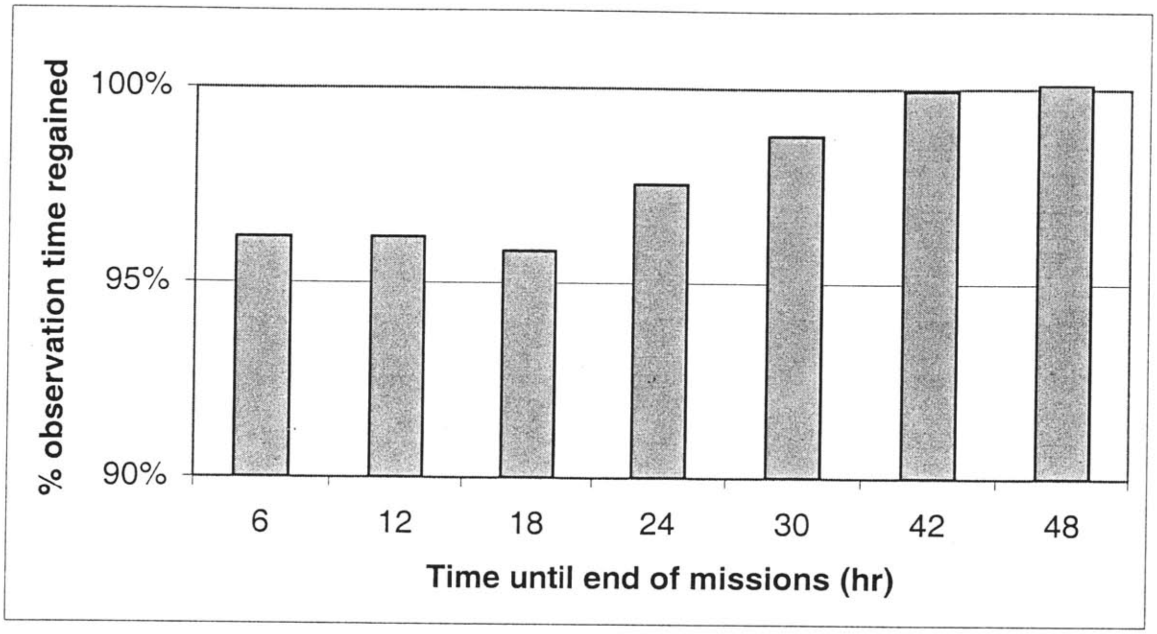


Fig. 6.3. Percent observation time regained by re-tasked satellite within specified time until end of mission.

Fig. 6.3 Discussion

This figure displays the percent of observation time that is regained by a re-tasked satellite from the amount of observation time that would have otherwise been lost from the original satellite due to the occurrence of an unexpected event. The percent observation time regained is shown as a function of the time remaining in the mission and is varied from 6 hours to 48 hours. The data presented in this figure is averaged over total mission times ranging from 1 day to 7 days, though no substantial difference was noticed between the percent observation time regained between the different mission lengths.

Fig. 6.3 Observations and Interpretation of Results

- The percent observation time regained increases as the time until the end of the mission increases. The percent observation time regained ranges from just over 95% of time regained to around 100% time regained. The only point that was observed to have a percent observation time regained less than around 95% was when less than 6 hours remained in the mission from the occurrence of an unexpected event. The percent time regained at these low times depended entirely on the satellite that was available, with some satellites able to perform additional observations and some satellites not able to perform a single observation. This will be discussed later in the chapter. After 42 hours approximately 100% of the observation time was regained. The exact amount is either a little more or less than what was originally planned for, as the new satellite that is re-tasked will have different windows of viewing opportunity from the original satellite. The practical implications of these numbers are that the integrated planner is able to regain a large majority of the total observation time that was lost due to the occurrence of an unexpected event at most any time during a mission, as long as it is not too close to the end of the mission. Near complete or complete observation time coverage is possible when an unexpected event does not occur within two days of the end of the mission.

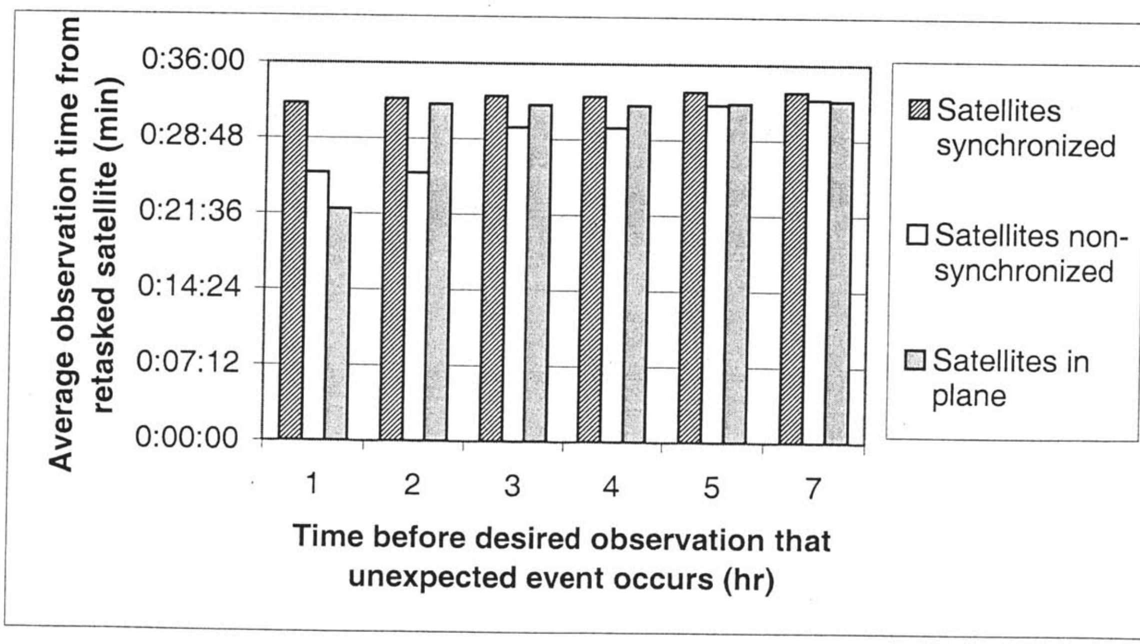


Fig. 6.4. Average observation time after failure for a 2-day mission.

Fig. 6.4 Discussion

This figure displays the average observation time that was achieved by a re-tasked satellite for the remainder of the mission when an unexpected event occurs at various times before an originally planned observation. Observation times achievable when unexpected events occur between 1 and 7 hours before the originally scheduled event are shown. As a point of clarification, the previous figures displayed observation time as a function of time remaining in the remainder of the mission. This figure and the following figures display the observation time that is achievable when an unexpected event occurs at different times before an observation that was originally scheduled but can now no longer occur because of the unexpected event.

The observation times are presented for three different types of satellites; satellites that are in synchronization with the satellite that is subjected to the unexpected event, satellites that are not in synchronization with the satellite subjected to the unexpected event and satellites that are in the same plane as the satellite subjected to the unexpected event. Satellites that are in synchronization with one another may be in different planes but will view the target at times that are close to one another. These satellites may be in synchronization for only a limited amount of time due to the nature of orbital mechanics and the satellites performing orbital maneuvers. Satellites that are not in synchronization with one another will have a spacing of several hours between respective viewing of the same target. This means that one satellite will view the target and the second satellite will view the same target several hours later. Satellites that are in plane (or close to being in plane with one another) either follow or lead one another. For low earth orbits, satellites that are in the same plane with one another will view the same target only minutes apart.

Fig. 6.4 Observations and Interpretation of Results

Comments that pertain to this figure have been included in the Observations and Interpretations section of the following figure.

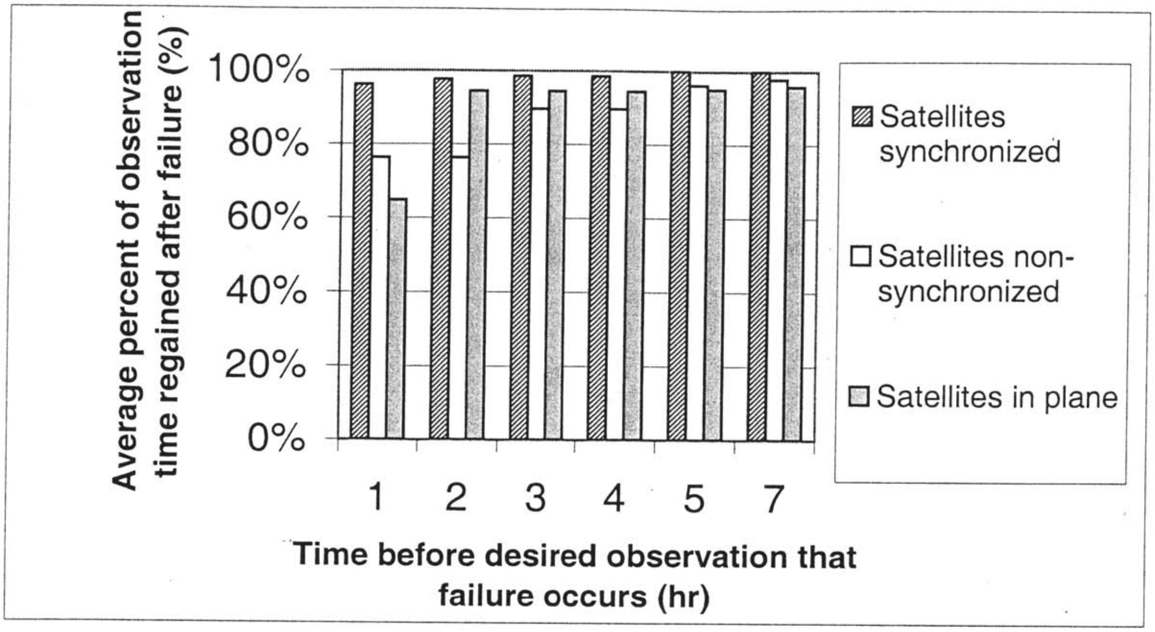


Fig. 6.5. Average percent of observation time regained after failure occurs.

Fig 6.5. Discussion

This figure displays the same information as the figure above, but as a percentage of observation time regained, as opposed to total observation time achieved. The trends in both figures are similar and will be discussed together.

Fig. 6.5. Observations and Interpretation of Results

- Satellites that are synchronized with the satellite subjected to an unexpected event are able to quickly regain most or all lost observation time. The percent of observation time that is regained is above 95% for all times and increases to near 100% if the unexpected event occurs around 5 hours prior to the originally planned observation. Because the satellites are in synchronization, little time is required before the re-tasked satellite is able to effectively replace the original satellite. To regain a large percentage of lost observation time with little warning, it is good to have satellites that are in synchronizations with one another. However, if the satellites are synchronized with one another, then all their observations will be tightly clustered. If it is desired that observations be evenly spaced out throughout the mission, then this is an extremely ineffective strategy for placing satellites. Also, if synchronization is desired, orbital maneuvers designed to increase the amount of time that a satellite can observe the target will destroy the synchronization between the satellites.

- Satellites that are out of synchronization with one another are able to regain a much lower percentage of the lost observation time than are satellites that are synchronized when the amount of time that an unexpected event occurs before a desired observation is low. For unexpected events occurring just one hour before a desired observation, the amount of observation time regained is near 80%. This is because the re-tasked satellite is not able to change its orbit in the amount of time given and will most often miss viewing the target near the time that was originally desired. The percentage of time that is regained increases with more time available to perform orbital maneuvers and approaches the 100% mark if the unexpected event occurs around 7 hours prior to the time of the original observation. For quick reaction times, satellites that are not in synchronization are not very effective at obtaining the initial desired observation. While the satellites can achieve most of the remaining observations that were originally going to be performed before the occurrence of the unexpected event, the first observation is almost always unattainable when the time is short. Satellites that are not in synchronization with one another do possess the trait that their observations are spaced out and not clustered.
- Satellites that are in plane or are nearly in plane with one another surprisingly have the poorest percentage of observation time regained when the time between the unexpected event and the desired observation time is short. After this time increases just a small amount, however, these satellites are able to immediately increase the percentage of observation time regained. The initial loss of observation time that was observed is attributed to the manner in which the EPOS 1.0 optimal planner has the satellites perform orbital maneuvers. Using the EPOS 1.0 optimal planner, satellites often have back to back viewing opportunities of the target. This means that the satellite performs an orbital maneuver that allows it to see the target on two or three sequential passes. If an unexpected event occurs before the first of these observations, a satellite that is nearly in plane with the original satellite will often not be able to position itself in an orbit that can also perform the sequential observations. So while the re-tasked satellite may be able to observe the target immediately in some

cases, it will likely loss the ability to observe the target multiple times in a row unless given an adequate amount of time.

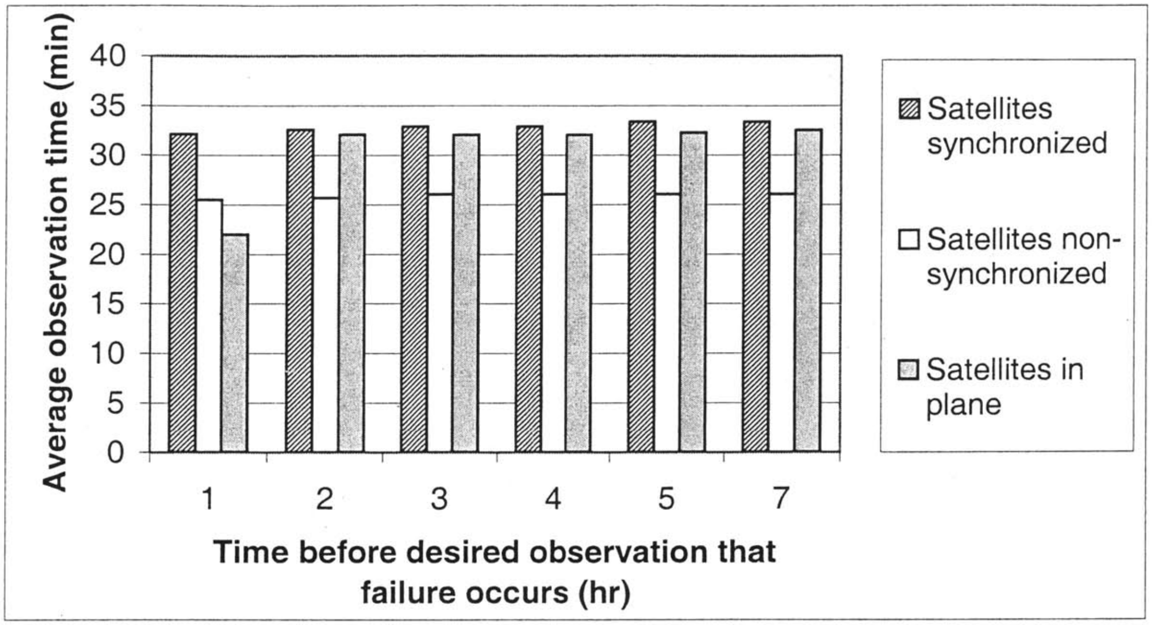


Fig. 6.6. Average observation time at or after desired observation time.

Fig. 6.6. Discussion

This figure displays that amount of observation time that is regained at or after the desired observation. This is different than the information displayed in the previous figure. The previous figures displayed all observation times that were achievable once the unexpected event occurred. This included some observations that occurred after the unexpected event occurred but before the desired observation would have occurred. This figure only presents the observations that occur after the desired observation would have occurred. Like the previous figure, this information is presented as a function of the amount of time before the desired observation that the unexpected event occurs. The results are also presented in terms of synchronized, non-synchronized and in plane satellites.

Fig 6.6. Observations and Interpretation of Results

Comments that pertain to this figure have been included in the Observations and Interpretations section of the following figure.

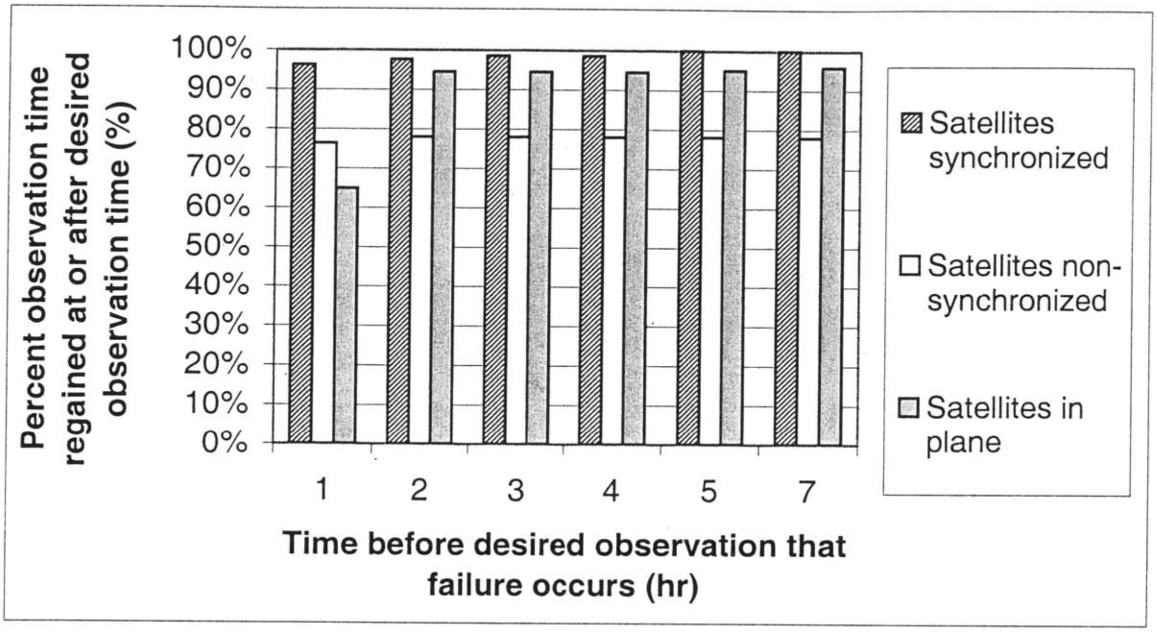


Fig. 6.7. Percent of time regained at or after desired observation time.

Fig. 6.7. Discussion

This figure displays the same information as the figure above, but as a percentage of observation time regained, as opposed to total observation time achieved. The trends in both figures are similar and will be discussed together.

Fig. 6.7. Observations and Interpretation of Results.

- The major difference between Figs. 6.6 and 6.7 and Figs. 6.4 and 6.5 is in the observation time regained with satellites that are not in synchronization with the satellite subjected to an unexpected event. As is seen in Fig. 6.7, the observation time that was regained by satellites that are out of synchronization is relatively constant for all times. Also seen in Fig. 6.7 is that the percent observation time regained is near 80%. Comparing this to the results seen in the previous figures, it is shown that the initial percent observation time regained when there is one hour between the unexpected event and the desired observation time increases when only looking at observation times that occur at or after the time that the desired observation would have occurred. Also comparing the remaining observation times achieved with the previous figures, it is shown that the percent observation times achieved are lower at longer times between the unexpected event and time of desired observation than they are in the previous figures. All of these can be explained by the fact that often the satellites that are not synchronized with the satellite experiencing the unexpected event will have an opportunity to observe the target before the desired viewing opportunity. Not counting the event that occurs before the desired observation would have decreases the percent observation time that can be regained. However, the amount of observation time that can then be achieved is very predictable. In practice, if an event is happening at the target at a specific time and there is no value or decreased value in observing the target before this time, then these are the figures that should be consulted. If all observations are important, then the previous set of figures have more bearing on determining which satellites can best regain the lost observation time.

Cost of Using Reaction Planner

This section presents results that show the cost of using the reaction planner in conjunction with the EPOS 1.0 optimal planner. Costs are presented in terms of additional fuel used for orbital maneuvers that are performed when using the reaction planner over that which would be necessary if an unexpected event occurred during the course of a mission planned using only the EPOS 1.0 optimal planner. The time during the mission when the use of the reaction planner is a viable option is examined as is the cost incurred as a function of the amount of time that elapses before a satellite is re-tasked and can view the target after an unexpected event occurs is also presented.

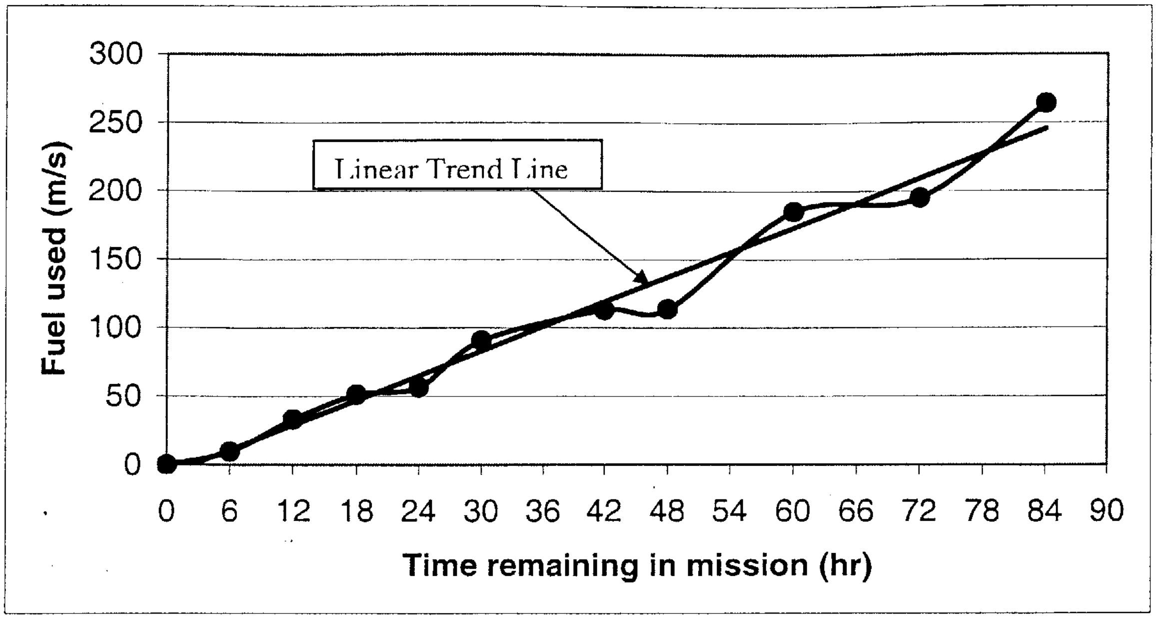


Fig. 6.8. Fuel used as function of time remaining in mission.

Fig. 6.8. Discussion

This figure displays the cost incurred by a satellite that has been re-tasked to view a target after the occurrence of an unexpected event. The cost that is incurred is presented as a function of the time that is left in the mission when the unexpected event occurs. Cost was observed for unexpected events occurring between 6 and 84 hours from the end of the mission.

Fig. 6.8. Observations and Interpretation of Results

- The cost incurred with the re-tasked satellite increases as the length of time that is available in the mission after the unexpected event occurs increases. This is expected, as there is more time for the satellite to perform additional orbital maneuvers to better observe the target when there is more time remaining in the mission.
- The cost trend observed in this figure is almost linear. This is expected as it was seen in the previous figures in the Benefits section that observation time increased in a linear fashion with increased time available until the end of the mission. In order to

achieve the increased observation time displayed, satellites will be required to continuously perform orbital maneuvers requiring an increasing amount of fuel.

- While the trend is close to linear, it is not exactly linear for the same reason as described in Fig. 6.2. That is, satellites will be performing a set of orbital maneuvers to better position the satellite to increase the observation time. This may result in performing multiple orbital maneuvers in sequence and then coasting for a set amount of time. This uneven performance of orbital maneuvers leads to a nearly linear trend that has discontinuities at certain times when maneuvers are being performed.
- The cost incurred as the time left in the mission approaches zero. This is because of the fuel cap placed on each of the satellites. When the time remaining for the satellite to perform maneuvers approaches zero, the size of any orbital maneuvers that a satellite would need to perform in order to observe the target become too large and burn too much fuel. This restriction means that the satellite will instead burn no fuel and as a result will not be able to view the target at all. At first glance, this may appear counterintuitive, as it is possible to imagine that the satellites would burn more fuel in the shorter amount of time they have available to squeeze in more observations. Squeezing in more observations in a shorter amount of time would mean that the satellites would need to make more drastic orbital maneuvers, which burn more fuel. This is not the case because the satellites are already able to maximize the amount of observation time that is possible by performing orbital maneuvers, so long as the fuel used does not exceed a predetermined limit. Because the satellites are already able to make orbital maneuvers to maximize the possible observations and because of the fuel limit, there is not a large increase in fuel used at shorter mission times. However, a closely related effect is observed and is discussed later in this section.

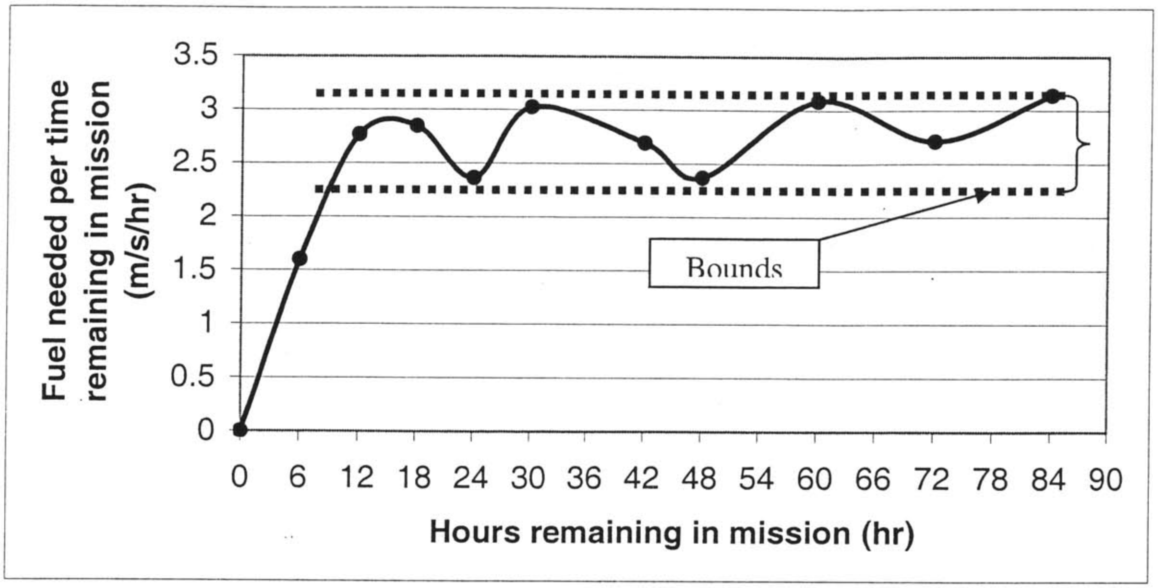


Fig. 6.9. Fuel needed per hour remaining in mission.

Fig. 6.9. Discussion

This figure displays the amount of fuel that is used to perform orbital maneuvers as a function of the time that remains in the mission after the occurrence of an unexpected event. The same range of mission time remaining after the occurrence of an unexpected event is examined in this figure as in the previous figure.

Fig. 6.10. Observations and Interpretation of Results

- This figure highlights some of the same trends discussed in the previous figure, namely the amount of fuel that is used for orbital maneuvers decreases as the time available nears zero. Looking at the 6 hours remaining data point, it can be seen from the above figure that not only do the satellites use a smaller absolute amount of fuel to perform maneuvers, the amount of fuel that is used on a per hour basis is also smaller.
- The remainder of the data shows that the amount of fuel that is used on a per hour basis is bounded from approximately 2.5 to 3 m/s being used every hour. This number is the amount of fuel that can be expected to be used by each satellite to obtain the maximum observation time possible, within the limits of total fuel used as specified. If additional fuel resources were available, the amount of fuel per hour used could be expected to increase. In practice, this number would need to be

determined before the mission began and would be based on the fuel resources available. Burning fuel at the rate of 2.5 – 3 m/s every hour can quickly become a large amount of fuel for mission times that would last days or weeks. Additional constraints on fuel use may be necessary to keep the satellites form continuously performing orbital maneuvers throughout the mission.

- As explained previously, because of the discontinuity in when the satellites perform orbital maneuvers, there are some irregularities in the numbers.

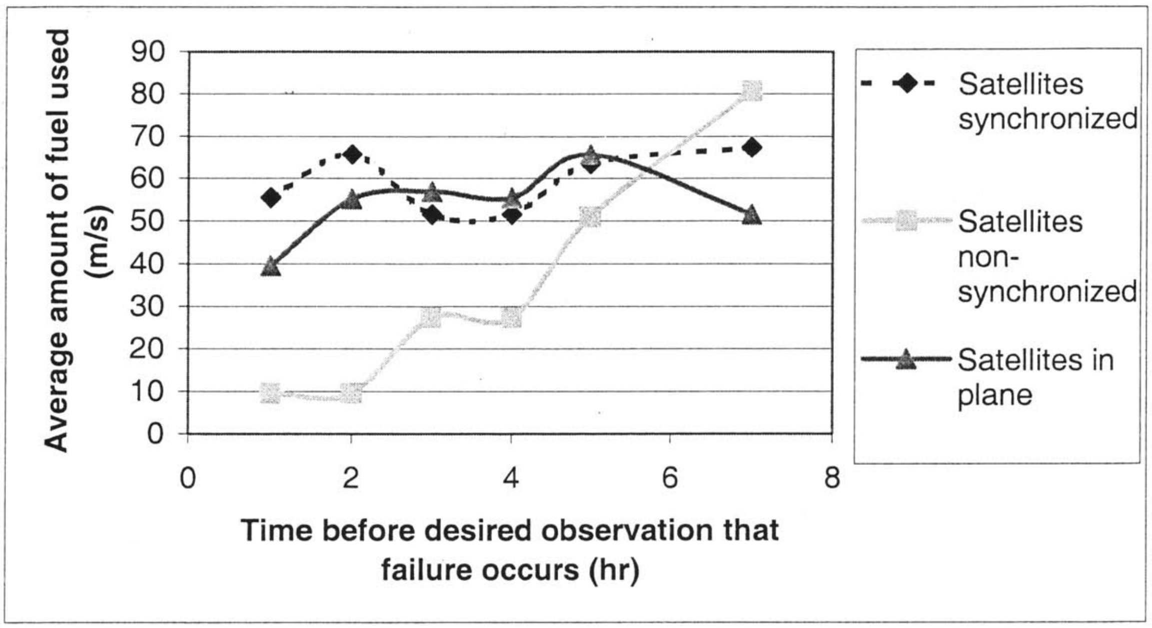


Fig. 6.10. Average fuel expended by re-tasked satellite.

Fig. 6.10. Discussion

This figure displays the average amount of fuel used by satellites as a function of time between occurrence of an unexpected event and the time of the desired observation. The distinction between satellites that are synchronized, non-synchronized and in a similar plane with the satellite subjected to the unexpected event are the same as those discussed for previous figures in the benefits section. The range of times between the unexpected event and the desired observation time is also the same as that used in the benefits section.

Fig. 6.10. Observations and Interpretation of Results

- The general trend of increased fuel usage as the time between the unexpected event and the desired observation increases is observed for all satellites, to varying degree. This is caused by satellites usually being able to make an additional observation if the time between the unexpected event and the desired observation is great enough. If this time is too short, then the satellite may not attempt the orbital maneuver. This trend holds to varying degrees for each type of satellite, as discussed in the following bullets.
- The fuel usage of satellites that are synchronized with the satellite subjected to an unexpected event is the most consistent across times of all the satellites. This is because the satellite is synchronized with the initial satellite and does not need to perform vastly different orbital maneuvers based on the amount of time given. Because the satellite starts out synchronized with the initial satellite, the amount of fuel that is used at small times is similar to that used at larger times.
- The absolute amount of fuel used for the synchronized satellites at small times is the largest of all the satellite types. While this may seem counter intuitive because these satellites are the closest to viewing the target, the relatively large amount of fuel being used on average is accounted for by the fact that most of the synchronized satellites will be able to view the target if an orbital maneuver is performed. Since all or most of the synchronized satellites are performing orbital maneuvers at small times, this keeps the average comparably higher than any of the other satellite classes.
- The fuel usage for satellites that are not synchronized with the satellite subjected to an unexpected event experience the greatest range of fuel usage of all the satellites. At small times, the fuel usage is the lowest, while at large times the fuel usage is the highest. The general trend of increased fuel usage is the same trend as previously discussed, which is that at small times most of the satellites that are out of phase will not be able to perform orbital maneuvers to view the target that meet the fuel requirement. As a result no maneuvers are performed by these satellites and the overall average fuel usage is small. As the time increases however, the satellite is able to perform orbital maneuvers that fall within the fuel requirements. Since the satellite is out of phase with the initial satellite a large burn must be performed to

allow the satellite to view the target. The size of the average burns performed for non-synchronized satellites is larger than burns from either of the other two satellite classes.

- Satellites that are nearly in plane with the initial satellite fall between the other two types of satellite classes in terms of fuel usage. At small times, only a fraction of the satellites in plane with the initial satellite will be able to perform orbital maneuvers capable of observing the target to the same extent as was initially desired. As the time increases however, more satellites can perform the orbital maneuvers to observe the target, increasing the average fuel usage. Since the satellites are nearly in plane with the initial satellite, the total amount of fuel that must be expended in performing orbital maneuvers is not as large as that needed for the satellites that are out of synchronization.

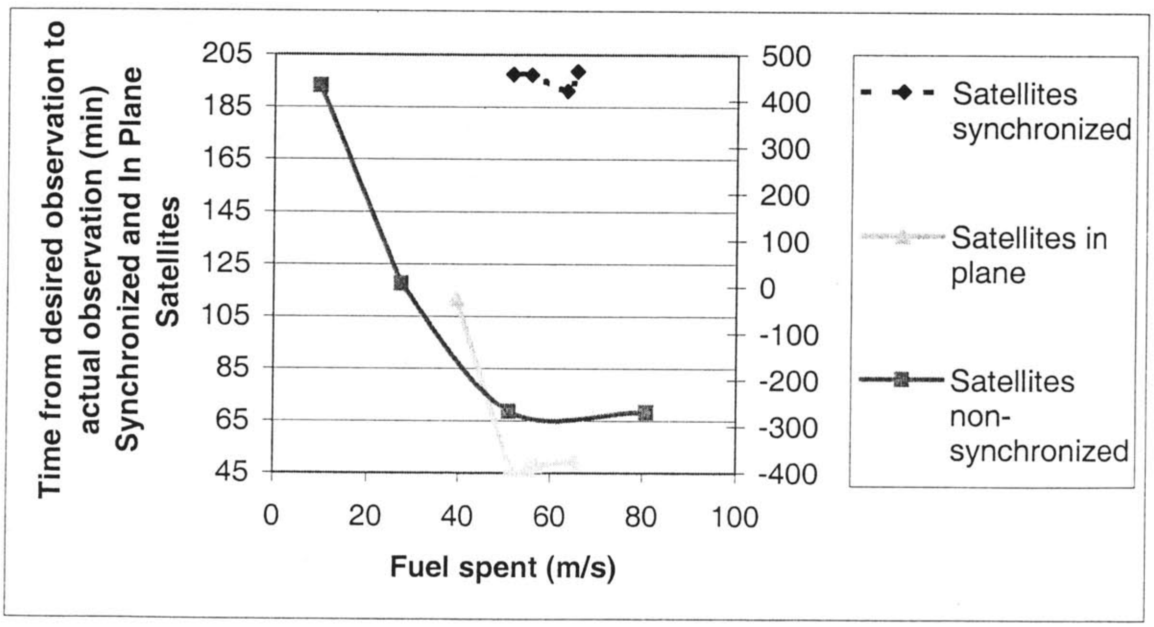


Fig. 6.11. Time from desired observation to actual observation as function of fuel spent.

Fig. 6.11 Discussion

This figure displays the amount of time that re-tasked satellites can begin to observe the target, measured from the desired observation time that was denied due to an unexpected event. The time that satellites in each of the three previously described classes take before performing an observation is displayed as a function of the amount of fuel used. Synchronized and in plane satellite times between actual and desired observation times are displayed on the vertical, left axis, while the non-synchronized satellite times between actual and desired observation times are displayed on the vertical, right axis.

Fig. 6.11. Observations and Interpretation of Results

- The general trend shown in all the satellites, with a few exceptions that will be discussed below, is that to achieve a lower time between an observation performed by the re-tasked satellite and the observation originally scheduled by the satellite affected by the unexpected event a greater amount of fuel must be used. This makes sense, as a premium will be paid to achieve an observation that is close time-wise to the originally planned observation. If the absolute time when the observation is to be performed is not critical, a smaller amount of fuel could be used by the satellite being re-tasked with the result being an observation that occurs later in time.
- The largest exception to the above trend is displayed in the synchronized satellites. It is seen from the data collected that there is a sharp upturn at the end of the data, which is also seen for the other satellites, but to a lower extent. For all the satellites this data is correlated with small times between the unexpected event occurring and the desired observation occurring. The attempt of the satellites to maximize the observations achievable overall in the time remaining in the mission dominates here. The result is that a large amount of fuel is used to squeeze in an additional observation, but that that observation may not occur near the desired observation. Since the satellites are all attempting to get the most observation time in the mission time remaining, and are not explicitly attempting to minimize the time between actual observations and desired observations, it is an indirect effect that this time is minimized, in general, by satellites using more fuel.
- The time measured on the vertical axis is between time from actual observation to desired observation. Satellites that are able to observe the target closest to the desired observation time are those that are near the plane of the satellite affected by the unexpected event. Synchronized satellites have the largest difference in time between

actual and desired observations. The satellites that are not in synchronization with the original satellite exhibit the largest range of values. These values range from several hours after the desired observation to a few hours before the desired observation.

- The range of fuel expended on achieving the observations also varies. Satellites that are synchronized and satellites that are nearly in plane require approximately the same amount of fuel and same range of fuel to perform the orbital maneuvers. The satellites that are out of synchronization have the largest range of fuel usage. The range is here an amount of less than one quarter that used for the other two satellites on the low side (and consequentially observation times over eight hours later than that desired) to an amount that is larger than ant of the other satellites (to achieve observations that occur before the desired observation time).
- If only observations that occur at or after the time when the desired observation would have occurred are desired, all the observations graphed for the synchronized and in plane satellites can still be used, while several of the observations that are possible would not be viable for the non-synchronized satellites. This is shown as all values of observation time difference below zero on the right, vertical axis (which measures the non-synchronized satellites).

Benefit/cost comparison

This section presents results that show the relationship between the benefit obtained and the costs incurred when using the reaction planner in conjunction with the EPOS 1.0 optimal planner. Benefits and costs are measured in the same manner as previously described. The benefit to cost ratio is examined as a function of time remaining in the mission as well as to the time with respect to the desired observation that the unexpected event occurred. The relationship between cost and benefit is also explicitly displayed.

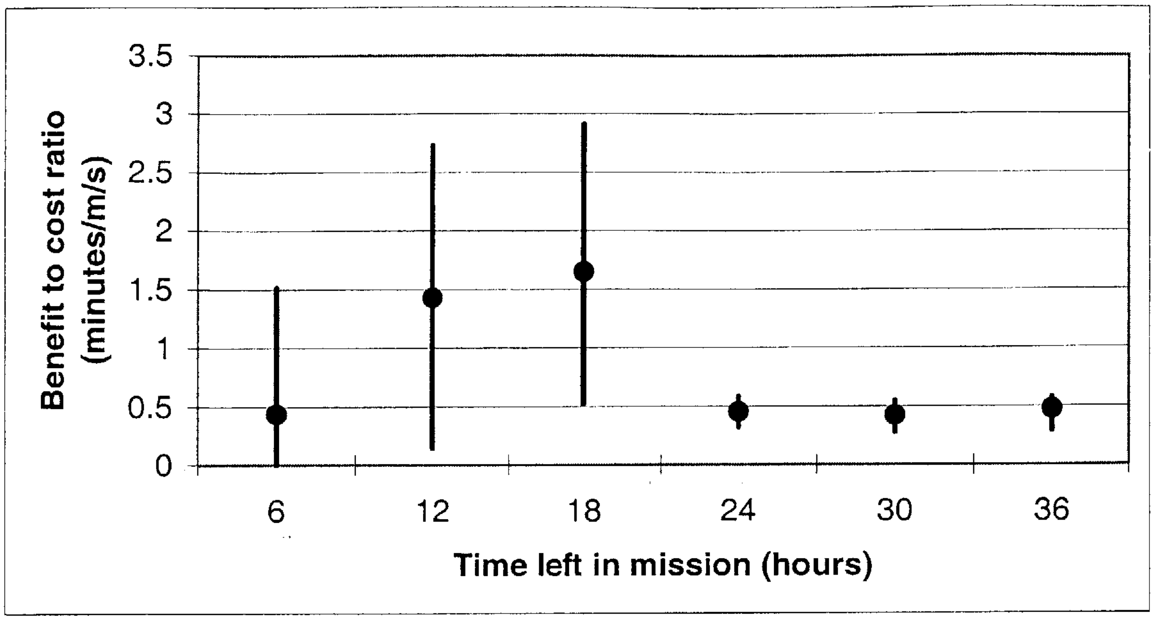


Fig. 6.12. Benefit to cost ratio as function of time left in mission.

Fig. 6.12. Discussion

This figure displays the average benefit to cost ratio that is experienced as a function of the time remaining in the mission after the occurrence of an unexpected event. The average values are represented by the data points, while the range of values are represented by the vertical lines through each data point. The same values as previously used for time until end of mission were repeated here.

Fig. 6.12. Observations and Interpretation of Results

- Most of the benefit to cost ratios fall near the 0.5 minutes per m/s of fuel used. The exception to this occurs near the 12 and 18-hour mark. This is for the reasons discussed above. That is, at low times, the satellites are not able to perform orbital maneuvers because of the expense involved. The result is that observation times are missed, lowering the overall benefit to cost ratio. At slightly higher times, the maneuvers can be performed and an extra set of one or more observations are obtained directly from these maneuvers. This extra increase in observation time results in a larger benefit to cost ratio. At even longer times, more maneuvers are made. This results in more observations, but the result is still not as large as the previous set of ratios.

- At shorter times remaining in the mission, the range of benefits to costs is large. This is because in the limited time possible, some satellites will be able effectively view the target either with orbital maneuvers or without orbital maneuvers and some satellites will not be able to see the target at all or will require much larger orbital maneuvers. At longer times, most or all the satellites can view the target by performing orbital maneuvers, which reduces the range of benefit to cost ratios.

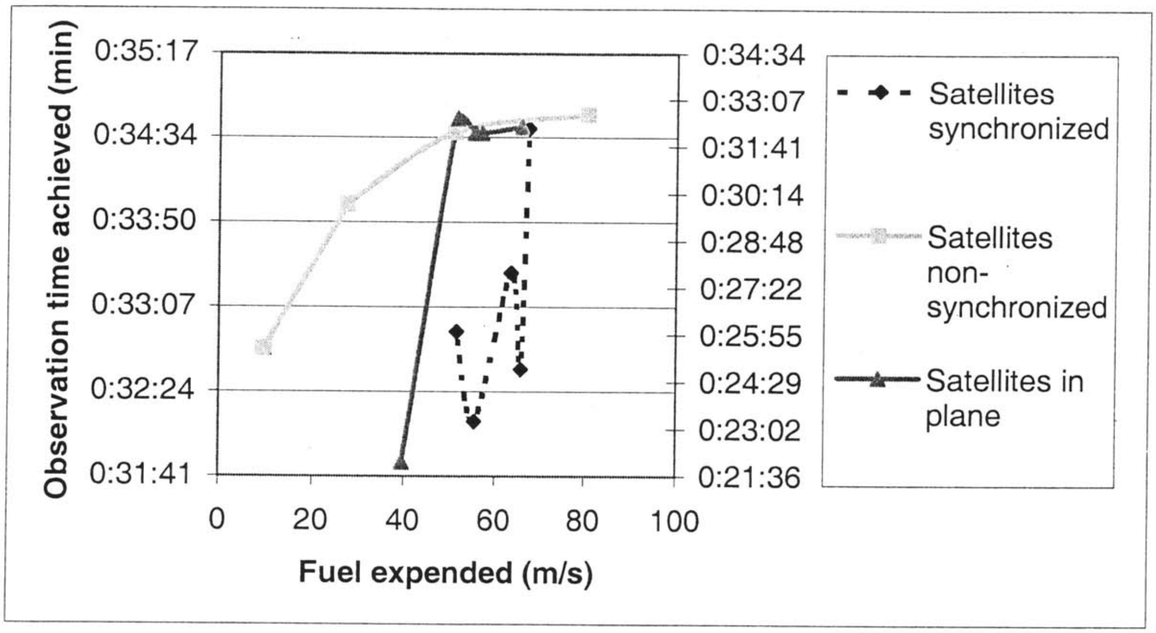


Fig. 6.13. Viewing time as function of fuel expended.

Fig. 6.13. Discussion

This figure shows the amount of observation time that can be achieved with the amount of fuel that is expended for the three classes of satellites previously discussed. The satellites that are synchronized with the satellite affected by the unexpected event have their observation time achievable measured on the vertical, left axis while the non-synchronized and in plane satellites have their achievable observation times measured on the vertical, right side.

Fig. 6.13. Observations and Interpretation of Results

- The general trend for all satellites is that increased observation time requires an increase in the amount of fuel expended while performing orbital maneuvers. A couple of exceptions to this are data points shown for the synchronized satellites. The reasons for this has previously been explained and relates to the use of large amounts of fuel to achieve observations that are achievable for a smaller amount of fuel later in the mission.
- For given amounts of fuel, the synchronized satellites have the best benefit to cost ratio of all the satellites. However, these satellites also require some of the greatest amounts of fuel to be used for all observations. The best benefit to cost ratio overall is achieved at the low fuel usage levels of the non-synchronized satellites.

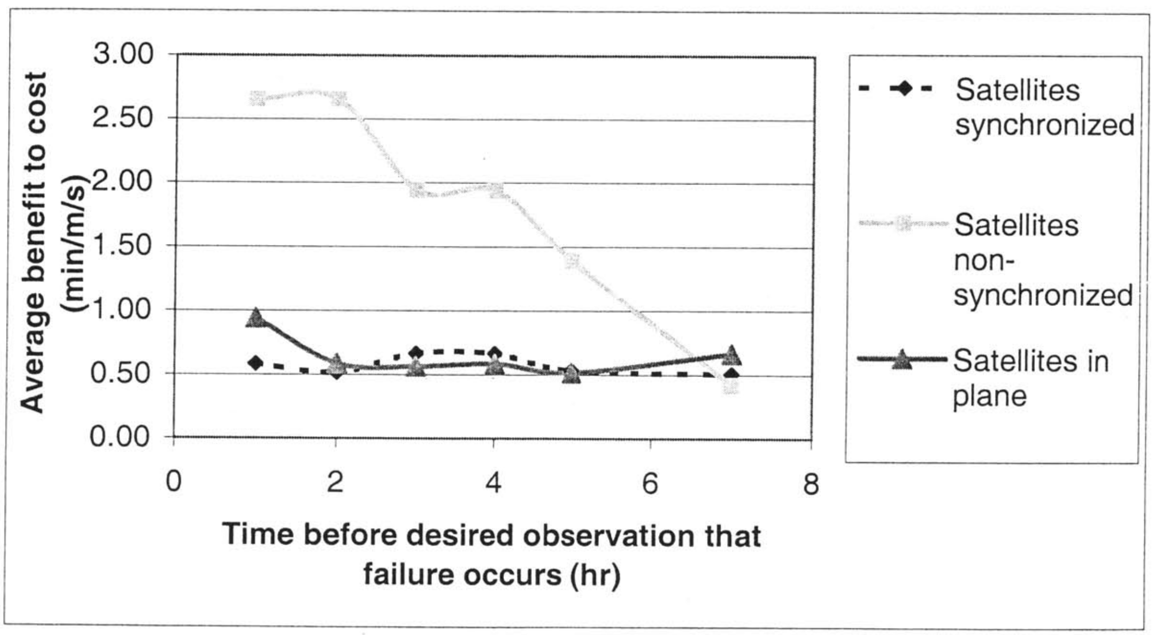


Fig. 6.14. Average benefit to cost of re-tasked satellite.

Fig. 6.14. Discussion

This figure displays the benefit to cost ratio for the three types of satellites previously described as a function of the amount of time before a desired observation time an unexpected event occurs. The range of times before the desired observation time that the unexpected event occurs is the same as previously used.

Fig. 6.14. Observations and Interpretation of Results

- The benefit to cost ratio for the synchronized and in plane satellites stays approximately constant for all times. There is a slight rise at the low times, due to the absence of orbital maneuvers for some satellites because of their large size due to the short time period.
- The non-synchronized satellites exhibit the largest range of benefit to cost rations, with the ratio being the largest at small times. This is due to most or all of the satellites forgoing orbital maneuvers at these times because of the large size of the maneuvers necessary. As the time increases, a larger proportion of satellites perform orbital maneuvers and the ratio approaches that of the other two satellite classes.

Guarantees from use of Reaction Planner

This section presents results that show the probability that the reaction planner will re-task a satellite to regain observations within a specified period of time. The time that a satellite achieves an observation from the time it is re-tasked is examined. This data is shown for synchronized, non-synchronized and in plane satellites, all of which have been previously described.

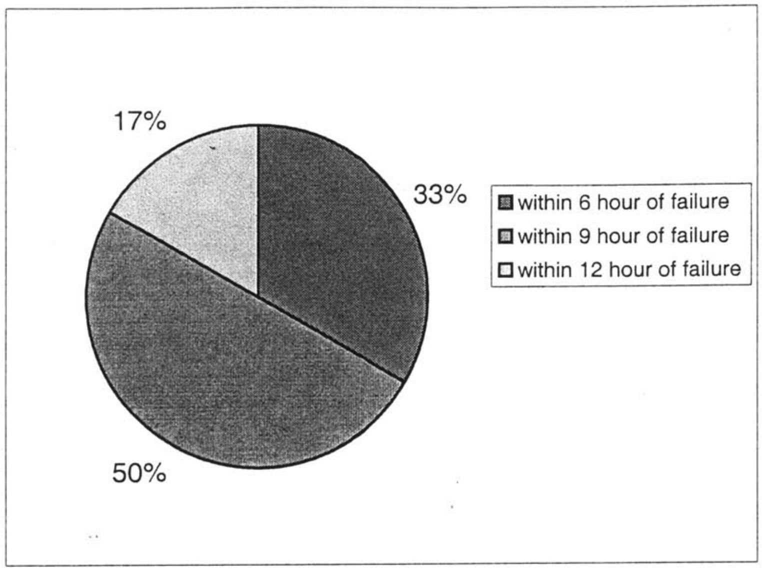


Fig. 6.15. Percent of Satellites Achieving an Observation within Specified Time of Failure (Synchronized).

Fig. 6.15. Discussion

This figure displays the probability that a synchronized satellite will be re-tasked and will achieve an observation of the target within 12 hours of the occurrence of an unexpected event.

Fig. 6.15 Observations and Interpretation of Results

Observations and interpretation of this figure is provided below under Fig. 6.17, due to the similarity of Figs. 6.15 – 6.17.

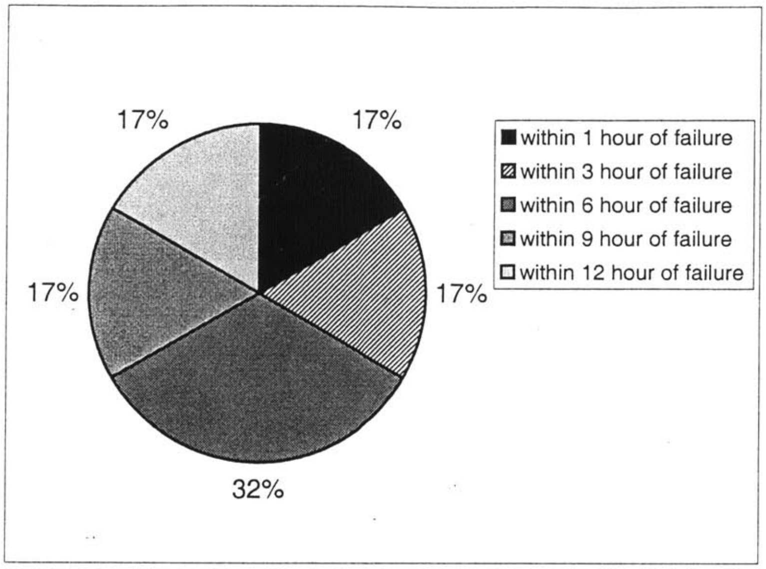


Fig. 6.16. Percent of Satellites Achieving an Observation Within a Specified Time of Failure (Non-Synchronized)

Fig. 6.16 Discussion

This figure displays the probability that a non-synchronized satellite will be re-tasked and will achieve an observation of the target within 12 hours of the occurrence of an unexpected event.

Fig. 6.16. Observations and Interpretation of Results

Observations and interpretation of this figure is provided below under Fig. 6.17, due to the similarity of Figs. 6.15 – 6.17.

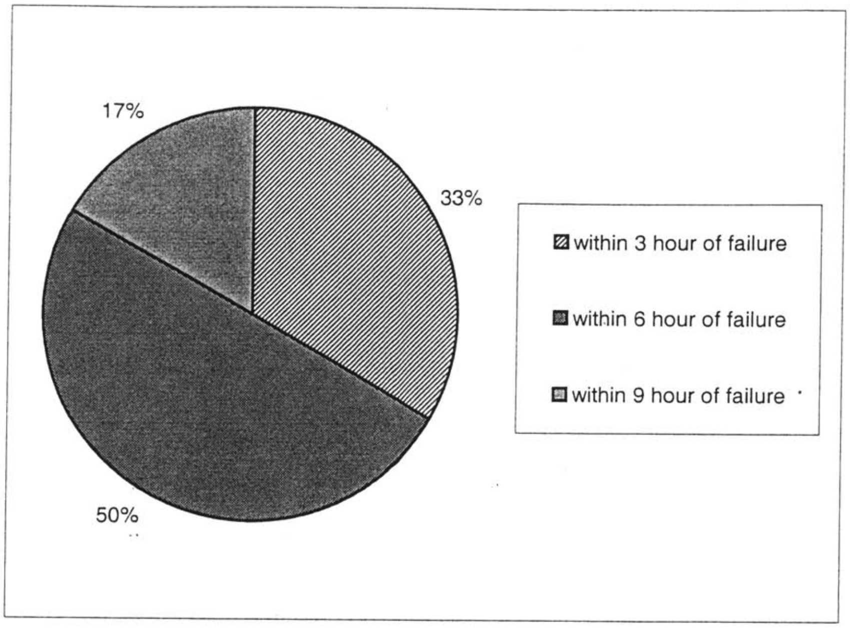


Fig. 6.17. Percent of Satellites Achieving an Observation Within a Specified Time of Failure (In Plane).

Fig. 6.17 Discussion

This figure displays the probability that an in plane satellite will be re-tasked and will achieve an observation of the target within 12 hours of the occurrence of an unexpected event.

Fig. 6.17. Observations and Interpretation of Results

- An observation can be achieved within 12 hours of the occurrence of an unexpected event for all satellite classes. This assumes that a working satellite is available to be re-tasked and that there is an ample amount of fuel available on the satellite.
- The in plane satellites have the greatest probability of achieving an observation within the shortest time of an unexpected event occurrence, on average. All in plane satellites are able to achieve an observation within 9 hours. However, observations within one hour may not be attainable.
- The non-synchronized satellites have the largest distribution of observation times of all the satellites. These satellites can either achieve the shortest time for achieving

and observation or the longest. The time required for these satellites to perform an observation is highly dependent on the actual satellite being considered for re-tasking.

- The synchronized satellites have, on average, the longest time required to observe the target. It is difficult for any satellite of this class to view the target in under 3 hours.

6.3 Summary of Results

Table 6.1 provides a summary of the results discussed above.

Table 6.1. Summary of results.

Parameter	General Trend	Synchronized Satellites	Non-Synchronized Satellites	In Plane Satellites
Benefit as function of time remaining in mission	General linear increase			
Percent observation time regained as function of time remaining in mission	Increases from around 95% to 100%. 0% at times less than 6 hours remaining			
Percent observation time regained as function of time before desired observation that unexpected event occurs		Highest average time regained, nearly constant, from around 95% to 100%	Lowest average time regained, low of approx. 75%, increases to approx. 95% at longer times	Second highest average time regained, low of approx. 65%, increases to approx. 95% at longer times
Percent observation time regained as function of time at or after planned occurrence of desired observation		Same as above	Nearly constant, just below 80%, loss of possible observation before desired event occurrence	Same as above
Fuel used	General linear increase			
Fuel used per time remaining in mission	Bounded for most values, low when time remaining in mission is low (due to high cost of orbital maneuvers for many satellites), no fuel used for very short remaining mission times (no maneuvers at all)			
Average fuel used		Largest amount of fuel used in most instances	Lowest amount of fuel used for low times and highest amount used for large times	Second largest amount of fuel used in most instances
Benefit to cost ratio	Nearly constant for most time, increase for short time but with large range			
Observations achievable	All satellites can observe target within 12 hours of unexpected event	Longest time on average to observe target	Largest range of times to view target, actual time depends on specific satellite	Lowest time on average to observe satellite

Chapter 7

Barriers to Satellite Pooling

This chapter is divided into sections discussing different barriers to satellite pooling. A section each for economic, political economy, organization behavior, legal and political barriers are presented below.

7.1 Economic Barriers

One of the primary reasons that the concept of pooling satellite resources between organizations has been considered as a concept, is to overcome economic barriers. Specifically, satellites are highly expensive resources to obtain and maintain. The number of satellites that are needed to provide complete and real-time coverage of multiple Earth based phenomenon may be beyond the resources of any one organization. Pooling satellite resources between organizations is one way in which the capital costs of satellites can be spread between multiple organizations, making the cost burden on any one organization realistic.

However, while the capital costs supported by any one organization are unrealistic, pooling satellite resources between organizations creates a new set of economic barriers that must be overcome. Three distinct economic barriers that exist due to pooling satellite resources have been identified. These are problems with: market and product definition, cost feasibility and cost distribution. Each barrier is discussed in more detail below.

Problems with Market and Product Definition

One of the first problems encountered with pooling satellite resources is related to problems with properly defining an appropriate market to serve. The customer base will vary greatly depending on what market is identified as appropriate. The definition of a market will also be necessary to define the product produced from the satellite system. The product is more than just information obtained from a set of satellites and includes delivery of the appropriate information in an appropriate format, to the appropriate customers, in a timely manner. The definition of the product and the market will drive the types of satellites that will be necessary to obtain the needed information, which will in turn drive how satellites are pooled together from various organizations.

Market Definition

The question of who could use the information provided is fundamental. Ideally, the satellites would be able to collect information on Earth based phenomenon and be distributed in a manner that could help save lives and minimize property damage. Under these idealized settings, people all over the world would be potential customers for such information, as over 100,000 lives are lost [15] and billions of dollars in property [31,33] are ruined annually.

However, to make this a reality will require more than just data collection from satellites. A sophisticated information distribution network will need to be provided to allow information that is collected to be processed and distributed to people in a timely and useful manner. It is assumed that some form of communication network needs to be in place that can reach the segments of populations that would most benefit from this information. This assumption raises the first problem with defining the market. Large segments of the developing world could be helped enormously with this type of information. For example, it is estimated that around 95% of the lives lost world wide due to natural disasters are in developing countries. [15] But the information distribution network that would allow the information to reach the appropriate people is not in place.

On the opposite end of the spectrum, most of the property damage that is sustained occurs in developed countries. In these countries, information distribution networks are much more highly developed, which makes it much easier to disseminate needed information to the appropriate people in a timely manner.

Also of importance is the ability to pay for the information. Developing countries and their population as a rule have a much lower ability to pay for this type of information than do developed countries and their populations. Similarly, most satellite resources are owned in developed countries. [42]

These factors influence how the market is defined. Should satellites be pooled to collect information on natural disasters occurring in developing countries that may cause great lose of life or should it be collected on disasters occurring in developed countries that may cause great property damage? If both, what are the priorities and in what ratio should the satellites be used to collect information on each? As each time a satellite spends time viewing a target, it is time that could be spent viewing a different target. This raises the question of who will pay for the information collected on each phenomenon? And should a satellite that has been paid for by one group of people be used to provide benefits to another group of people, which is essentially satellite subsidies?

These types of questions must be answered when defining the market that satellite pools will service. How the market is defined will either create or remove other barriers, more of which will be discussed later in this chapter. As satellite pooling uses resources from several different organizations, it is anticipated that many different viewpoints on what is the proper market to service will be raised. Enough of these viewpoints must be adequately accommodated so that a suitable number of satellite resources will be available.

Product Definition

Closely related to the problems associated with defining a suitable market, is the problem of defining a suitable product. The market definition will in effect define what the product must look like. For example, if it is determined that more accurate hurricane landfall predictions is the information that should be provided to residents of Florida and states with a coastline on the Gulf of Mexico, the types of satellites needed and the frequency at which they are needed will be defined based on this market definition. Not all satellites will be capable of collecting the types of information that would be most useful for this task. Finding which organizations have satellites that could accomplish the technical requirements of observing hurricanes is essential. Also critical will be ensuring that an organization that owns this type of satellite is amenable to the way in which the market is defined. For example, a satellite owned by the Russian Federation and capable of providing the needed information on hurricanes may not be available to the pool if the only benefit that the satellite will provide is to citizens of the United States. The issues of market definition and determining how organizations with satellites will be compensated for the use of their satellites is critical to the product definition, which is in turn critical to making a satellite pool work.

Cost Feasibility

While pooling satellite resources is designed to help mitigate the costs that any one organization would have to bear in providing real-time coverage on a variety of Earth based phenomenon, pooling creates costs as well. Pooling is designed to minimize the capital costs associated with satellite procurement and deployment that organizations would have to bear. These additional costs act as a barrier for enacting a pooling system. Costs incurred because of pooling include: coordination costs between organizations and loss of satellite utility. Both of these are discussed below.

Coordination Costs

Direct monetary costs are just one part of creating any satellite system, especially a dynamic system composed of a set of satellites pooled from various organizations for a short time period. There are several costs that are incurred by all organizations participating in a satellite pooling arrangement that relate to the coordination of the satellite pool. These include time and effort costs, and political capital costs.

Currently, designing, deploying and operating a satellite is relatively routine but still highly complex. Coordination must be maintained between various stakeholders, such as satellite users, designers, operators and financers. Once in orbit, the satellite's usage is strictly and tightly prescribed, often with elaborate and technically complex scheduling and planning algorithms that seek to optimize the satellite's utility [25]. Planning the satellite's operation requires coordination between users and operators and is usually discussed and planned in meticulous detail, far in advance of the actual execution of the plans.

Creating a pool of satellites that are used in a temporary manner to observe phenomenon in real-time, often with little advanced warning, would require that the pooled satellites be available at a moments notice to organization controlling the pool. This is completely different than the way in which satellites are currently controlled. The sudden transfer of satellite control to another organization with little notification is not typical, or even seen as desirable. To ensure that the satellite's primary mission is not completely lost, coordination between the satellites' primary organization and the pooling organization must occur. This requires great time and effort to set up channels of communication, ensure that the proper and needed information flows both ways between organizations and guarantee that the satellites are properly returned to the primary organization. This is time and effort that organizations would have to spend to make satellite polling feasible, beyond what they need to expend normally. Resistance to costs incurred due to time and effort is a barrier to making pooling feasible.

Political capital costs are also a part of the costs associated with coordinating organizations to facilitate pooling. Currently, organizations control their satellite relatively autonomously. Even satellites that are financed by the U.S. federal government are all operated independently of one another in different programs. And even within programs, for example NASA, each subdivision operates autonomously and with their own objectives and procedures. Coordinating all these organizations together and aligning their interests to the extent necessary to participate in a pooling system will require that a great deal of political capital be expended, either in convincing the primary organizations directly to participate or forcing their cooperation by invoking a higher authority. It is envisioned that political capital costs will be greatest at the start of the creation of a pooling system and that once organizations become embedded and familiar with the pooling system, less political capital will need to be expended to maintain participation.

Loss of Satellite Utility

Satellites are currently designed with a specific mission and life span. The range of satellite costs can range over multiple orders of magnitude [5] and the utility of satellite actions may be difficult to measure [16]. From the point of view of the primary satellite organizations, any loss of satellite time is viewed as a loss of satellite utility, as satellite time is usually completely filled. Further, actively using the satellite to engage in burns to transfer between orbits may shorten the life of the satellite. The loss of satellite utility and the shortening of operational life span is not offset by the cost to procure or operate the satellite. Instead, the operational cost of the satellite may increase due to the additional complexity and uncertainty involved with participating in the pool. The result is that the efficiency and effectiveness of the satellite, as measured by the original satellite mission goals, will decrease for all satellite participating in the satellite pool. This type of systematic decrease in system performance is not desirable to primary organizations, as they will have a harder time satisfying users and financiers, which could also affect future satellite procurement decisions. This systematic loss of efficiency and effectiveness will likely make primary organizations resist any attempt to include their satellites into a pooling arrangement.

Cost Distribution

The cost distribution of a satellite pooling system is of critical importance. Primary satellite organizations design and deploy their satellites to service a specific need for a specific set of users. Primary organizations then finance the satellites based on this need and user base. Government organizations such as NASA or the NRO submit a budget that includes a request to fund a satellite for a specific mission, which ultimately has to be approved by Congress. These organizations do not have the authority to unilaterally change the defined mission of their satellites after they have received funding. A similar situation occurs in university satellites, which are often funded by government sources to accomplish a specific purpose. Satellites belonging to private industry are financed to serve a particular market segment. While private industry has some flexibility to re-define the mission that their satellites will accomplish, they are also bound by financing and any contractual agreements that they enter into to provide services from their satellites.

As each organization finances the satellite for a specific mission and user base, changing the mission to service a different set of users raises the question of who should pay for this loss of satellite time that the primary organizations incur. Further, the costs associated with using the satellite pool to observe Earth based phenomenon encompass more than just compensating the primary organizations for their loss of satellite time. Costs include operation costs for the pooling organization, development of an information network to distribute information and costs associated with the development of an infrastructure to service satellite on-orbit.

Issues of equity and feasibility over who pays for this service are paramount. The manner in which these issues are addressed will determine who should finance the operation of the pooling system.

7.2 Political Economy Barriers

Several barriers exist at the intersection between politics and economics that act to inhibit the formation of satellite pools. Though not always obvious, the failure of market and political institutions acts as formidable barriers. For the specific case of pooling satellite resources, organizations have multiple market and political driven motivations to resist formation and participation in a satellite pooling system. Different types of organizations will exhibit different types of resistance. The types of market and political failures causing this resistance have been identified as: the presence of unstable property rights, collective action dilemma, and coordination problems. The following section elaborates on each of these and identifies primary organizations that may be susceptible to these failures.

Unstable Property Rights

One of the primary categories of problems associated with satellite pooling is the presence of unstable property rights. By design, the satellite pooling organization does not own any, or at least very few, of the satellites that it needs to accomplish its mission. Also by design, the primary organizations utilize their satellites the majority of the time for whatever purpose fulfills its mission objectives. The result is a large set of satellites, all of which have a dual mission. Each is designed for its primary mission, which the satellite spends most of its time executing. Each satellite in the pool also has a secondary mission, which is being available for use in the pool to observe Earth based phenomenon in real time. The unstable property rights problems stems from this dual mission and the shifting ownership of the satellite between organizations, each of which has very different needs. Three specific unstable property right type problems caused by satellite pooling have been identified: ownership and responsibility issues, problems associated with provision of public goods and a tragedy of the commons problem. Each are discussed below.

Ownership and Responsibility Issues

What organization actually owns the satellite and when raises several issues. For instance, does the primary organization own the satellite at all times and just temporarily lends or leases it to the pooling organization? If so, which organization has responsibility for the satellite during the time that the satellite is participating in the pool? Which organization is liable if the satellite experiences a failure when in the custody of the other organization?

The uncertain status of the satellite ownership and responsibilities of each organization creates a barrier for organizations to become involved in a pool as it exposes their organization to additional liability and uncertainty than they would otherwise be exposed to. Two examples illustrate this point. First, if a satellite fails while under the operation of the primary organization due to the operation of the satellite from the pooling organization, it potentially loses the utility of the satellite for its own users. Second, if a satellite fails while under the control of the pooling organization because of circumstances that are the responsibility of the primary organization and because of the failure the satellite pool is not able to warn against an imminent disaster, the primary organization opens itself up to liability.

Sharing the operational responsibility of the satellite but not the ownership creates uncertainty for which organization is liable for accidents, failures or misuse. As primary organizations would prefer to avoid this added uncertainty, it creates a barrier for their participation in satellite pooling activities.

Provision of Public Goods

As the satellites that are available for use in the pool has some of the attributes of a public good from the view point of the pooling organization, the problem of investing in the satellites becomes a problem. The pooling organization has low incentive to spend any of its resources in providing for upgrades or maintenance for any of the satellites that it uses. As there are ideally many satellites available for use in the pool, the pooling organization may not have the resources to spend on maintenance or upgrade on all the satellites. The pooling organization has little incentive to spend its resources on upgrades that would help better accomplish its mission, as it utilizes any satellite only a fraction of the time it is in orbit. The pooling organization also has little incentive to spend resources on satellite maintenance as there are other satellites that may be better maintained currently available. As the pooling organization utilizes any satellite for such a short amount of time there is a high incentive to let primary organizations invest its own resources, as it utilizes the satellite for the vast majority of the time and is the only satellite available that can accomplish its mission. It would also probably be obvious to most primary organizations, if not immediately then by precedent, that the pooling organization has little incentive to spend resources on satellite maintenance or upgrades. This poses a barrier to primary organization participation in a pool, as their organization is essentially being asked to subsidize a different mission in addition to their own primary mission.

Similarly, a potential moral hazard exists in the use of satellites by the pooling organization. As there are many satellites available for use in the pool, there is little penalty on the pooling organization for misuse of the satellites. If a satellite is misused, even to the point of damaging relations with a primary organization, the pooling organization still has many other satellites to choose from. The pool design of not relying on any one organization to provide all the satellites creates a sort of satellite insurance. The absence of any one satellite or primary organization will likely not be detrimental to the pooling organization. This creates the potential for a moral hazard, from the pooling organization not being as careful as they should with the satellite resources entrusted to them from the primary organizations. The possibility that the pooling organization may not have the incentive to ensure safe operation and return of satellites is a large barrier to primary organizations desiring to participate in the pooling system.

Tragedy of the Commons

A tragedy of the commons problem occurs when a public resource that is freely available for use is overused to the point where it is no longer a use to anyone. A modified version of this problem is believed to be a potential barrier to satellite pooling. In the traditional tragedy of the commons problem, several actors all over-utilize a public resource until the resource is of no use to any of the actors. The reason for this over-use is the common belief that if personal use is not made of the public good, then some other actor will use the good anyways.

In the satellite pooling problem, only the pooling organization makes use of the satellites, which from their point of view is a public good supplied by the primary organizations. However, the satellites that are available for use are not homogenous. Rather, some satellites will be better suited to fulfill a particular objective than other satellites. This may be due to the type of sensors on the satellites, the age of the satellite or the satellites initial orbit. For any reason, it may arise that a sub-set of the satellites available to the pooling organization is seen as more desirable than other available satellites. If the satellites with the highest utility to the pooling organization are consistently used at great frequency, the utility of these satellites to the primary organization will dramatically decrease. As the decrease would be due to over use by the pooling organization, the primary organization would have incentive to either prohibit their satellites usage or mask the utility of their satellite to the pooling organization, making it appear more undesirable for use. The end result is either that fewer primary organizations will be willing to participate in a pooling system for fear of over use of their satellites or that true satellite capabilities available to the pool are hidden from the pooling organization.

Collective Action Dilemma

The creation of a satellite pooling system is also susceptible to classic collective action dilemmas. Specifically, over or under representation of stakeholders in the pooling organization will be a barrier to creating an effective pooling system and various freeriding problems are also barriers. Both of these are discussed in more detail below.

Over and Under Representation

Adequate representation of all stakeholders is essential to creating an organization that is effective. Classical problems of over and under representation are potential problems in creating a pooling organization. Over and under representation usually occur when strong, concentrated interests receive more representation than diffuse interests. For example, large multi-national corporations are better able to marshal their resources and effectively lobby for their interests. Environmental groups however have a much more difficult time garnering support for their cause, even though the environmental group ideally represents the interests of a much larger group of people. The environmental group has a more difficult time being effective because their support base consists of such a large group of small stakeholders, where the mobilization of all these people is difficult to attain.

The problems of over and under representation are anticipated to be a problem for the formation of any pooling organization as well. As the pooling organization will obtain most of its satellites from organizations that have satellites available, the interests of these organizations will likely dominate the mission and mode of operation of the pooling organization. Stakeholders that have a large interest in information that the pooling organization could provide, but are unable to offer either satellite, fiscal or other resources, may have a much smaller voice in the formation and operation of the pooling organization.

As a large proportion of the world's satellites are registered as U.S. satellites [42], it stands to reason that U.S. satellites would form a large segment of the satellites used by a pooling organization. Many developing countries have either very few or no satellites at all, but could still benefit greatly from the information provided from the pooling organization. Although there are many more people represented by developing countries with few resources, their interests will likely be under represented when forming a pooling organization, when compared to U.S. interests.

Freeriding Problems

The problem of freeriding, where one group receives benefits while another group bears the costs, is another barrier to forming a pooling organization. There are several freeriding problems associated with the formation of the pooling organization.

As discussed previously, the problem of the pooling organization engaging in freeriding with the primary organizations' satellite resources is a potential problem. The pooling organization has incentive to benefit from the investment that the primary organizations have made, without making a substantial investment in satellite resources itself. From the viewpoint of the primary organizations, this creates a substantial freeriding problem with the pooling organization and will act as a barrier to primary organizations participating in the pooling system.

There is also freeriding among users of the information from the pooling organization. It is anticipated that the pooling organization will receive funding from some source, either public or private or a mix. As the information that is collected must be distributed widely and quickly for it to be effective, the information will likely also be available to groups that have not help financed the pooling organization at all. Two examples illustrate this point. First, it is assumed that information collected will be distributed with priority to people or countries that have helped create the pooling organization. If a natural disaster affects a large region, perhaps territory in multiple countries, it is possible that people will receive the same information regardless of whether their nation has contributed to the pooling organization, because by necessity the information will need to be widely distributed. Second, private corporations can also act as freeriders. Additional information on natural disasters would greatly help profitability in some industries, such as insurance companies. It would be very easy for entire industries to make use of the additional information provided without helping to support the pooling organization.

Coordination Problem

Technically coordinating all the satellites that will be needed to gather the required information on Earth based phenomenon will be a large challenge for the pooling organization. This is because all the satellites and primary organizations have their own unique operating system and procedures that are in place. There is little or no compatibility between these systems and procedures. Either the pooling organization will need to adapt to adapt to all the individual systems and procedures used by the primary organizations or some new standards will need to be adopted by the primary organizations.

Solving the problem of technical coordination creates other non-technical problems. Technically, it may not be feasible for the pooling organization to accommodate all the different procedures that are in use. This leads to the necessity that primary organizations adopt some standardization to allow the pooling organization to more efficiently utilize the satellite resources. The type and scope of this standardization across primary organizations is likely to meet stiff resistance, especially if the new standards are intrusive and differ greatly from procedures currently in place. Any necessity to adopt new standards that do not support the primary organizations' primary mission will likely result in resistance of participating in a pooling system.

If standardization does occur across the primary organizations, even to some degree, the pooling organization will likely utilize this standardization as a basis to achieve greater efficiency in operating the satellites. While standardization does offer greater efficiency gains, it also grants a competitive advantage on the pooling organization, who will likely help create and then utilize the new standards. The effect will be that the pooling organization may gain a competitive advantage over any other potential future competitors. If the practice of utilizing satellite resources from different organizations became commonplace, the emergence of a market offering pooling services may emerge. If it did emerge, it may be difficult to dislodge the initial pooling organization and create a competitive market.

7.3 Organizational Barriers

Currently, organizations with satellites have sole discretion in deciding how the satellites should be used. Additionally, the same organizations have a large, if not dominant, role in creating the satellite's mission and design. This control over all aspects of the satellite life cycle has lead organizations to expect and take for granted their ability to control the satellites. Any change, by sharing control with a pooling system, will encounter resistance along structural, cultural and political lines within primary organizations. Resistance along each of these organizational lines will be discussed below.

Structural Resistance

Structure is the officially designed aspect of an organization. "Structure" is used to describe an organization's mission, management hierarchy, how it is divided into subunits, etc. Primary organizations are structured in manner that gives them authority over most aspects of their satellites. The primary organization controls in large part what type of mission their satellites should accomplish (i.e., space research, weather observation, military reconnaissance, communications, etc.), how the satellite should be designed, securing funding for the satellite, and operation of the satellite. Primary organizations are designed to be relatively self contained and autonomous in terms of designing, deploying and operating satellite systems. Being autonomous, primary organizations are not used to sharing control of their satellites resources. Additionally, as the primary organization mainly controls the design process, the satellites are optimized to accomplish the goals of the primary organization, as opposed to the goals of the pooling organization. The control that primary organizations currently have over their satellites and the manner in which satellites are designed creates structural barriers within organizations that will make it difficult to participate in a pooling system. Specific aspects of the structural barriers are: divergent needs and ability to influence the design process of satellites and missions, lack of established organizational links between primary and pooling organizations, and integrating organizational capabilities to satisfy two very different user bases. Each of these is discussed below in more detail.

Divergent Needs and Abilities to Influence Satellite and Mission Design Process

One of the primary responsibilities that organizations have is to ensure that the products and services that they offer adequately fulfill the needs of their users. To accomplish this, organizations spend a great deal of time working with the users of their satellites when designing the satellite mission, hardware and operational plan. Working closely with the users through the design phase and into the operation of the satellites is essential in ensuring that the satellites offer the required performance. Working closely together, organizations become, hopefully, attuned and responsive to the needs of their user base.

Lack of Established Links between Primary and Pooling Organizations

Currently, primary organizations exercise almost autonomous control over their satellites. The concept of sharing control with another organization and the interfaces required to do so are non-existent. The lack of linkages between the organizations creates problems in satellite operations, mission prioritization, an unclear hierarchy of satellite control and problems in accounting for the satellites.

Satellite operations are currently handled exclusively through one organization, sometimes the organization owning the satellite handles operations and sometimes it is a different organization that has been contracted to fulfill these responsibilities. The process of repeatedly handing off operational control of a satellite, as called for in a pooling system, has never been done. There is no organizational structure in place that facilitates an interface between different organizations that will each control the satellite for different periods of time.

Additionally, the manner in which operational control of satellites is currently managed is not conducive for the type of control transfer that is necessary for a pooling system. One of the largest expenses in the life cycle costs of a satellite are the operational costs. These costs come primarily from maintaining a ground support center, staffed with professionals that are always on call or are present in the operational control center. It would be highly undesirable from the point of view of the primary organizations to have the ground support personnel on site when the satellite is unexpectedly transferred to the poling organization, or to have no ground support personnel on site if the satellite is suddenly transferred back to the primary organization. The direct costs and the extra risk involved with either having unnecessary or too few ground support people on hand are high.

Mission prioritization is currently determined within a primary organization. The resultant set of mission priorities are usually the result of long, iterative negotiations between members in the organization and the user base and are usually agreed upon well in advance of the actual mission operations. Participation in the pooling system changes the process of prioritizing mission objectives. If a satellite is temporarily unavailable to a primary organization because it is being used in a pool, the satellite will have missed the opportunity to fulfill some set of objectives that the primary organization has previously defined. Once returned to the primary organization, a decision must be made on how to prioritize the rest of the mission. Namely, has the lost opportunity been completely lost or should it be rescheduled? And in the new circumstances that the satellite is in, have the remaining objectives kept the same priority, or do those need to be re-evaluated? The re-evaluation of mission priorities and capabilities on such a compressed time scale is usually not desirable and usually has no established precedent. Additionally, the negotiations that carefully crafted the original mission priorities between the different stakeholders may not have time to be re-created, giving some group more control over the new priorities than they originally held.

The command structure that controls a satellite is usually well established in primary organizations. Ultimate authority for the satellite is known at all times. Using the satellite in a pooling organization makes the determination of who has control and authority over the satellite more difficult to determine. While it is relatively easy to determine who has authority when the satellite is under the control of either the primary or the pooling organization, at the interfaces of this control, ultimate command authority is more difficult to determine. For example, while the satellite is under the control of the primary organization, who has the authority to re-task the satellite to participate in the pool? Does the primary organization have the power to override this decision? Who has the power to make these decisions when confronted with non-standard operating conditions?

The question of who has ultimate authority over the satellite is related to which organization has responsibility for the satellite's use or misuse. Any incident that leads to a satellite failure or some other operational failure, such as failure to provide information, occupation of a wrong orbit, interference with other satellite systems, could have serious consequences for the organization with authority over the satellite. As there are serious consequences, the organization with responsibility will also want to ensure that they have control over the satellite's operations also.

Currently, all costs associated with a satellite's operation are covered by the primary organization. When a satellite participates in a pool system, two different organizations are using the satellite system. What share of the operational costs and what share of the benefits derived from the satellite usage is a non-trivial question. For example, while the pooling organization uses the satellite only a fraction of the time of the primary organization, the pooling organization uses the satellite in a much more intensive and high risk manner. Specifically, engaging in a series of orbital maneuvers and servicing rendezvous is a much more risky and life shortening set of activities than simply letting the satellite coast in its original orbit. What level of additional risk and stress the satellites are subjected to should determine the level of costs that each organization should bear. Determining the levels of risk, stress and assigning a cost to each is non-trivial. Additionally, if the pooling organization gains income in some manner from the information that it provides, is the primary organization entitled to a portion of these proceeds, as it was their satellite that was used to obtain the information?

Satisfying Different User Bases

The base of users for the satellite services offered by the primary organizations and the pooling organization are vastly different. The user base of the pooling organization is interested in timely information on Earth based phenomenon that poses an immediate and great danger to life and property. The user bases of individual primary organizations vary greatly from the user base of the pooling organization and from one another. Primary organizations user bases could include a variety of scientists, meteorologists, the intelligence and military communities, satellite cell phone customers, etc. It is expected that while each user base will be sensitive to the needs of the other user bases, that this sensitivity will not translate into support for sacrificing their needs to satisfy the needs of a different user base. As organizational structures are designed in part to service the needs of their user base, conflicting user base needs will lead to conflicting organizational goals.

The desire of an organization to fulfill the needs to their traditional user base may lead to several conflicts between primary and pooling organizations. Conflicts over the amount that missions will be compromised, equity issues between user bases, misalignments of incentives between organizations and misrepresentation of satellite abilities to hide satellites from the pool are all possible. The attempt to satisfy different user bases could lead to organizational barriers making it difficult for primary organizations to participate effectively in a pool system

Cultural Barriers

The culture of an organization has as great an impact on how the organization operates as the structure does. Each primary organization has the history of being self sufficient, with the need for self-sufficiency reinforced by a culture developed to accommodate this need. By participating in a satellite pool system, the culture of autonomy that is prevalent in all primary organizations would be challenged. And trying to change the culture of an organization is difficult, as cultural habits are often ingrained and happen without conscious effort.

The participation in a pooling system will mean a significant change in the way that primary organizations view their relationship with the satellite. Currently, the satellite is under the organization's control at all times. With a pooling system, there will be some times when the satellite is not under the control of the primary organization. While these times may not be long in duration, the uncertainty in timing and frequency of loss of control will pose a different degree of challenge to different organizations. While the severity of the loss of control will be driven in part by the requirements that the satellite is fulfilling for the primary organization, the openness to the sharing of authority with another organization is also critical in determining if a primary organization would be willing to participate in a pooling system.

Culturally, some organizations are very closed or rigid, and may have a hard time operating in an environment where their satellites may sometimes pass to another organization. Military organizations, which have a culture of closely supervised control, may find it difficult to participate in a different culture that would allow certain satellites to pass to another organization. Other organizations, such as research or academic institutions, are created to explore new avenues of generating knowledge. Culturally, these types of institutions would be more open to "trying something new". Commercial institutions have a culture that is driven by ensuring that the profit seeking motive of the organization is fulfilled and to reduce risk of previous investments. If participation in the pooling system is designed to adequately compensate organizations for the increased uncertainty and risk, commercial organizations would be willing to participate.

Another cultural pitfall that could be a problem is the "We've never done that before here" or "We did not come up with the idea" mentality. If an organization is loathe to try something new because it was either not developed in house or because it has never been attempted before, it will be difficult to elicit participation in the pooling system.

Political Barriers

The political processes within and between organizations is a large barrier towards primary organization participation in a pooling system. Political barriers are often erected when changes in organizations are proposed. Change within organizations often will result in the loss of power, prestige, influence, etc. of one or more stakeholders, who will then resist the proposed changes to preserve their interests. The sharing of satellite resources with a pooling system could easily be interpreted as a loss of power over the primary organization's satellite resources. Political resistance from within the organization and between organizations can be expected to resist inclusion in a pooling system.

Political Barriers within Organizations

Resistance from groups within primary organizations that have a vested interest in maintaining complete internal control over satellite resources will be expected to offer resistance to inclusion in a pooling system. Groups could include user groups, which would have their mission objectives occasionally subsumed by those set by the pooling organization, and satellite operators, which would have to relinquish control to the authority of the pooling organization. Both these groups would be expected to resist because of the loss of power and loss of influence in setting satellite mission priority. These groups would stand to lose power and influence both to the pooling organization and to new groups established within the primary organization that would be tasked with interfacing with the pooling organization.

Political Barriers between Organizations

Resistance to participation in a pooling system can also be expected from organizations. Having their satellites subsumed within a pooling system threatens the authority, autonomy and relevancy of primary organizations. Primary organizations may fear that additional power could be lost to the pooling organizations or that their participation in the pool threatens their support base for obtaining future support, funding or credit.

7.4 Legal Barriers

Several key legal barriers exist in the creation of a pooling system. These issues pertain to both the changing status of control of the satellites in the pooling system and the availability of information collected through the system. The primary legal barriers identified were; organizational legal liability, availability of information collected from the pool system, intellectual property issues and cost of the distribution of information.

Organizational Legal Liability

The question of which organization has legal liability for the satellites is a potential barrier to a pooling system. The temporary transfer of satellite control from the primary organizations to the pooling organization and back again, complicates this question. While some aspects of the legality problem are relatively easy to workout, such as mis-operation resulting in satellite failure, others are not.

One of the difficult legality issues concerns satellite life span. Most satellites in use today require and are capable of relatively little orbital maneuvering. Conversely, for use in the pooling systems, each satellite must be capable of performing repeated orbital maneuvers. Once in a space environment, one of the highest sources of stress on the satellite comes from orbital maneuvering. If use in the pooling system subjects a satellite to substantially additional stress than what it was designed for, it could impact the life span of the satellite, shortening it below what was expected. Liability for satellites that experience failure significantly before their design life expires is difficult to determine, as the pooling organization, primary satellite operator and satellite manufacturer may all be held liable, but be able to point fingers at other organizations. Similarly, if a satellite fails while under the control of the pooling organization, but no mis-use was found, is the pooling organization liable?

Another difficult liability assignment question concerns the primary organization opening itself up to liability from the general public. It is not difficult to envision a scenario where a satellite tasked by the pooling system fails and as a result adequate information is not able to be obtained about a deadly and imminent Earth based phenomenon, the result of which is a lose of life or property. The primary organization could open itself up legally for not maintaining the satellite in an appropriate condition to fulfill the terms of its agreement to lease satellites for use in the pooling system. Similarly, if an Earth based phenomenon does not make "the cut" as something worth observing by the pooling system and it results in deaths that could have been avoided if observed, both the pooling and primary organizations could be found liable.

Availability of Information Collected from Pooling System

Ideally, satellites in the pooling system could be used to help respond to Earth based phenomenon occurring anywhere in the world by providing information that could be useful for a response. However, this ability requires the collection and dissemination of information. Sometimes more information than was intended to be collected can be obtained and sometimes this capability could be misused. Occasionally, collection of this type of information is prohibited.

When collecting information on Earth based phenomenon, it is easy to see how more information than is required could be collected from a certain geographical region. This is easy to envision if the type of data needed is in the visual or infrared spectrum for a small geographic area, where the information returned is for the full footprint of the satellite sensor. If other activities are occurring nearby, for example military exercises or a military campaign, it is conceivable that information about troop movement and capabilities could be unwittingly collected and disseminated. This could pose a potential risk if hostile forces were to "call in" a faked phenomenon that they requested satellite coverage to help respond to, when in actuality they were looking for this "extra information".

To prevent this type of use of satellite imagery during wartime, the United States has in the past prevented imagery from being sold by commercial satellites over a particular geographic region. The most recent example of this was during the military involvement in Afghanistan. Here, the US government indirectly prevented the selling of satellite imagery that could compromise its forces by entering into an exclusive contract through the National Imagery and Mapping Agency with commercial satellite imagery suppliers for all imagery over Afghanistan for the duration of the conflict. [9, 11]

The problem with this type of satellite "shutter control" is if there was a genuine threat that could be countered with the use of the pooling system and the dissemination of the information was hampered due to security restrictions. In addition to the legal issues discussed above that this raises, this is in direct violation of international space law. In the United Nations' Principles Related to Remote Sensing of the Earth from Outer Space, Principles XI and XII explicitly talk about the sharing of satellite imagery data that could help avert natural disasters in affected states [22].

Another concern is the prohibition of collecting data from satellites at all over some regions and countries. For example, a natural disaster occurring in the Middle East, specifically near Israel, would be difficult for a pooling organization to respond to. This is because US law states in the 1996 Kyl-Bingaman Amendment to the National Defense Authorization Act for Fiscal Year 1997 that collection of detailed satellite imagery relating to Israel is prohibited [28]. Legal prohibitions over selected areas like this make response to legitimate disasters legally impossible in some areas.

Intellectual Property Issues

One of the most contentious legal issues concerned with data collected from satellites is the issue of intellectual property (IP). In the United States, the precedent has been established that data collected and returned from government satellites is open to the public and is available to any user at the marginal cost of the information. And data collected from commercial satellites is protected under IP rules, allowing commercial firms to restrict access to the information that their organization has collected and to sell it in limited form for a profit. This IP framework is not the same standard used throughout the world, specifically in Europe. European IP laws essentially protect information collected from government satellites, raising the legal question of whether satellite data is a public or private good [24,9]. This effectively limits commercial exploitation and use of this information. This is important for a pooling organization that may include satellite from several different countries. This is even more important if the satellite organization is specifically designed to encourage commercial usage of the data returned from the satellites, to try and increase the beneficial usage of the pooling system.

Distribution of Information

Related to the question of commercial usage of public satellite resources, is current US government policy. OMB (Office of Management and Budget) Circular A-130 states that environmental data obtained from Earth observing satellites should be distributed to the public at the marginal cost [24]. Any pooling system that incorporates commercial entities must be careful to distinguish between information availability to the general public and opportunities for commercialization of that information.

7.5 Political Barriers

Two key political barriers were identified that could impede the development and deployment of a pooling system. These political barriers are anticipated to come from politicians and satellite users.

Politicians

One of the most intense political barriers is expected to come from politicians. Two reasons for this resistance have been identified. First, politicians are often hesitant to allow resources that they have funded for a specific purpose to be used for a different purpose, especially if the new purpose will decrease the value of the resource in terms of fulfilling the original purpose. This "re-tasking" of satellite resources is in effect creating new legislative priorities by shifting resources between missions. Politicians are expected to resist this tendency that the pooling system will have to "re-legislate". Secondly, politicians that represent organizations that may lose some of their satellite resources to the pooling organization are likely to fight for the exclusion of their constituent from the satellite pool.

Users

Satellite users who fought to help get funding for satellites to meet their needs are expected to resist any pooling system initiate that would decrease the amount of time that "their" satellites are available for use. This is expected, as satellite resources are always in short supply and are hotly contested over. Additionally, there is a dislike for the "free-riding" from the pooling organization. Users who expended political capital to fund satellites that meet their needs would be loathe to see another organization reap the benefits of their hard labor effectively for free.

Chapter 8

Creation of a Pooling Organization

This chapter contains sections discussing the needs of a pooling organization capable of overcoming the barriers discussed in the preceding chapter. Sections discussing the needs of the pooling organization and the categories of stakeholders with interest in a pooling organization are first presented. Next, several types of organizational models that could be used to base the pooling organization on are identified and discussed. Each of these models is then evaluated for applicability to the problem of overcoming the identified barriers, and one organizational model is downselected. Additional information is presented for the design of the pooling organization, followed by an overview of benefits and drawbacks to using the organizational model identified.

8.1 Pooling Organization Needs

In the past, data collection on Earth based phenomenon via satellite has been the exclusive domain of government agencies. [29] However, as satellite technology has become more advanced, as people have found new uses for information collected by satellites and more information is collected, commercial and not-for-profit organizations have also moved into Earth observation with satellites.

Technically, this growing interest has come from the availability of data in increased amounts and at increased resolutions, often at a quantity and quality once the sole domain of intelligence agencies.[6] In addition to the quality, the quantity of data is also rapidly increasing. It is estimated that over the next decade, data available from NASA satellites will provide a hundred-fold increase in the amount of data returned from satellites. [3] Politically, this increase in data is threatening to change the mission of government agencies involved in satellite observations of the Earth. For instance, in the area of Earth observation, NASA's mission is morphing from an agency with a mission to provide experimental satellites and improved data sets to being forced to manage and disseminate a quickly growing set of data.

As the data returned from Earth observing satellites increases and the uses for that data becomes more sophisticated, there is also a growing need to integrate data from different satellites into one central database[7]. [3] Integrating this data will need long-term

[6] The end of the Cold War saw a radical increase in the quality of satellite imagery available for commercial and scientific use. In the early 1990's, in an effort to raise hard currency, Russia started offering satellite imagery taken with previously top secret satellites and offering a resolution of 2m. To keep from being left behind, the United States in 1992 passed the Land Remote Sensing Policy Act to encourage domestic provision of commercial satellite imagery. [sat 5, 22] Another example is in astrophysics data with the National Virtual Observatory project's Earth Observing System Data and Information System (EOSDIS) Core System (ECS). This data base is supposed to provide "one stop shopping" for multidisciplinary data in a timely manner. A next generation database management system called SEEDS (Strategic Evolution of ECS Data Systems) is planned that will offer; 1)data processing, 2) value added product generation such as data mining and format translation, 3) , 3)

[7] There has been some effort at responding to this need for an integrated database. An example of actions taken to integrate data is Earth Science Information partners. This partnership has been operating for the last three years and is an attempt to integrate all Earth observation data collected from different NASA funded organizations. Another example is in astrophysics data with the National Virtual Observatory project's Earth Observing System Data and Information System (EOSDIS) Core System (ECS). This database is supposed to provide "one stop shopping" for multidisciplinary data in a timely manner. A next generation database management system called SEEDS (Strategic Evolution of ECS Data Systems) is planned that will offer; 1) data processing, 2) value added product generation such as data mining and format translation, 3) archiving and distribution of data, and 4) service to help users find and acquire data [sat 2].

consistency in satellite data collection [24] and cross-institutional cooperation, both of which are outside the scope of current organizational operations.

These new technical and political forces indicate that there is a need to carefully consider the type of organizational structure that a pooling organization will take on. To decrease the costs of running and participating in a pooling system, the pooling system needs to be multi-functional. This means that the data that is collected must appeal to a broad range of users, likely from the research, commercial, weather and emergency response communities, which are divided between the government, private, academic and not-for-profit sectors.

The following section briefly identifies different general stakeholders, followed by a discussion of what types of organizational structure a pooling organization could be modeled after.

8.2 Pooling System Stakeholder Categories

Five general categories of stakeholders were identified as being critical when determining the type of structure that a pooling organization should be modeled after. Each stakeholder and their general needs are briefly described below.

Research Scientists and Academic Community

To successfully perform research and understand Earth phenomenon, researchers need full and open access to data. Additionally, as most researchers have access to limited funding, the data provided must be available at a relatively low cost, preferably no higher than marginal costs. The goals of the research and academic community are to advance their research and provide education through access to scientific data.

Industry

The overriding objective of industry is to obtain, use or disseminate data in a manner that provides a reasonable rate of return on the company's investment and to try and ensure that there will be a stable, future market potential for the goods and services that they provide. The most common model that industry uses to make a profit is by restricting access to data that they either generate of manipulate in some manner. It is not in industries interest, nor is it viable, to provide data at marginal costs when the initial capital outlay for satellite procurement must be recouped. While industry is a large user of "free" data provided to the public domain from government sources, industry would also prefer that they do not have competition from government agencies that provide this free data. The needs of the industrial sector vary somewhat between data providers and companies that offer a product or service by adding value to information that others provide.

Government Agencies

Government agencies have an overriding mission to use public funding in a manner that protects life and increases economic value of property. Government agencies involved in Earth observation usually do this either through collection and dissemination of data or through the development and testing of new sensors and technologies. Generally, the data that is collected is either used by the research community or used to maintain the capabilities of the weather forecasting infrastructure. To increase the value of the data that is collected and disseminated, industry is often urged to use the data to create new commercial products, increasing the value of the data returned with government satellites. Government agencies have also recently acted as brokers between the scientific community and commercial Earth data suppliers, providing such services as data calibration, data verification, intellectual property negotiations, price targets and delivery schedules. [24,6,28]

Policy Makers

Policy makers generally want to ensure that their constituents' needs are well supported. Constituents, of course, can have a widely varying set of needs. Domestically, policy makers support both the diffuse needs of the general public as well as the more concentrated needs of a few satellite and satellite data related corporations. Both sets of constituents have a vested interested in a pooling system and the data it could provide, but for sometimes different reasons. Internationally, there is a strong desire for an increase in information that could be provided to countries, especially developing countries, that would help mitigate devastating effects of natural and man-made disasters, both of which the pooling system could help with. The aims of policy makers concerned with domestic or international use of the pooling system differ somewhat in responding to different constituents' needs.

General Public

The general public is best served by the provision of the best information possible, at an acceptable cost. This information will only be of value to the general public if it is widely distributed and processed into a form that is both usable and delivered in a timely manner.

8.3 Pooling System Organizational Models

There are several choices for the type of organizational structure that the pooling organization can be modeled on. These range from existing government agencies, new consortiums of existing agencies, new government agencies, academic institutions, not-for-profit institutions, private companies, and private-public partnerships. Each of these organizational types is discussed below, along with advantages and disadvantages associated with each that would be relevant when forming the pooling organization.

Existing Government Agencies

One of the most likely choices for taking on the responsibilities for the pooling system would be within an existing government agency. Two types of existing government agencies could be potential candidates for hosting a pooling system. These are agencies with research driven missions and operational missions. Each is discussed below.

Research Agencies

Government agencies like NASA have as their main mission the advancement of science, technology and knowledge. In pursuit of this mission these agencies, with NASA specifically in mind, have engaged in large programs that are both technically and institutionally complex. Examples include the Apollo program, the Space Shuttle program and the International Space Station. The last two in particular are relevant to the choice of an agency such as NASA to create and operate the pooling system, as both are large, complex and long-lasting programs. These programs demonstrate that the technical ability for creating and operating a pooling system is within these organizations. However, it is questionable whether the long-term operation of a pooling system would fit with a research driven mission. Even the Space Shuttle program, which is the closet NASA has come to maintaining a long-term space program, is not operated by NASA. Rather it is operated by the United Space Alliance, a consortium between Lockheed Martin and Boeing. While NASA does provide mission control personnel for long-term satellite missions, these missions are usually not designed to need continuous human interaction outside of a relatively short timeframe and in no cases is needed at a high level on an ongoing basis. This lack of experience and lack of mission focus on operational issues makes research agencies like NASA a poor choice for operating a pooling system.

In designing and implementing a pooling system, while NASA has the technical ability, there would be the potential of a conflict in interests. As the pooling system would pull satellites from several different government agencies, having one agency determine how its' and other agencies' satellite resources will be used could result in political clashing between the organizations. The political aspect of this makes it unlikely that an organization like NASA could effectively create and implement a pooling system.

Operational Agencies

Government agencies like NOAA have as their primary mission the provision of a continuous service that provides benefit to the public. In pursuit of this mission, these agencies attempt to determine what services are needed to be effective. NOAA and the National Weather Service in particular provide weather and climate information to the research community and to the general public. This is accomplished through the procurement, operation and management of a relatively large fleet of Earth observing satellites. However, while these agencies have research programs, many of the radical technologies fielded by these agencies were initially developed from other agencies like NASA. This lack of deep experience or expertise in large scale research may make the scheduling, planning and data management problems associated with designing and implementing a pooling system beyond the reach of operational agencies.

In providing for the operation and management of large technical systems, agencies like NOAA have years of experience. However, the real-time data acquisition mission focus of a pooling system is different than the current mission focus of short-term weather observations. This change, or addition, of mission focus is difficult to accomplish well in an organization, without being shortchanged by the culture built around the current mission. As an example, when NOAA's mission was enlarged to include data acquisition for long-term climate change, there was much debate on how well NOAA was actually carrying out this new mission. A similar concern would be had for incorporating the pooling system into an existing organization's mission. Additionally, expanding the scope of one agencies mission by using resources from a different organization may prove to be politically difficult, if not impossible.

Coalition of Government Agencies

Instead of placing responsibility for the pooling system in the hands of one organization, another alternative would be to form a coalition of existing government agencies and place responsibilities there. A previous example in the area of Earth observing satellites of this taking place is with the NPOESS (National Polar-orbiting Operational Environmental Satellite System). The NPOESS is a relatively newly formed coalition between NOAA, Department of Defense and NASA to bring all weather and climate observing satellites together under one integrated, civilian system. This system is run with personnel donated from all three agencies. In 2005, the system will expand to include a European weather satellite (MetOp) [36]. Operationally, the existence of NPOESS proves that a coalition of existing government agencies could be formed to operate and manage a pooling system. However, it took 8 attempts over 25 years before civilian and military weather satellites were actually integrated together into one system. So while possible, it is very difficult to actually form a coalition between existing agencies.

Formation of a New Government Agency

An entirely new government agency could be created for the sole purpose of creating and operating the pooling system. This would be in the precedent of NASA's creation for space exploration and NOAA's creation for weather observation. A new agency could be constructed to possess the required technical and operational experience necessary to create and operate a pooling system. The greatest obstacle that would inhibit the creation of a new government agency that would have jurisdiction over the satellite resources of other government agencies would be political. Political resistance from existing agencies could kill the creation of any new agency at conception. Additionally, the ideology against the expansion of government is also relatively strong, making creation of a new agency a difficult political sell to law makers as well.

Academic Institutions

Academic institutions have expanded upon their traditional role of involvement in basic and applied research and have entered into the realm of spacecraft design, deployment and operations. Organizations like the Harvard-Smithsonian Center for Astrophysics (CfA) have expanded beyond providing scientific knowledge and experiments for spacecraft. The CfA designed, built, and operates the center for the Chandra Observatory, one of NASA's "Great Observatories", as well as provides data analysis from information collected from Chandra. Drawing from top research universities and research laboratories from around the nation, academic institutions have the technical skill to create a pooling system. The reputation of academic institutions as being unbiased and above the political process makes academic institutions a potential choice for being an organization that could effectively create a pooling system. However, the relative size of these institutions compared with other possible organizations is relatively small. Additionally, current political trends have demonstrated that government funding for these institutions is being shifted to other agencies, such as the National Science Foundation. This lack of size may make it difficult to create a coalition powerful enough to gather support from primary organizations that would have to contribute their own satellite resources.

Not-For-Profit Organizations

Not-for-profit organizations like the SETI (Search for Extra Terrestrial Life) Institute hold much of the same appeal as academic institutions for creating and operating a pooling system. Namely, that the lack of affiliation to any government or private entity gives them an aura of credibility that an equitable pooling system could be created. Top non-profits also usually have no problem attracting top talent, meaning that it is possible that the technical resources could be obtained. However, many non-profits, like SETI, draw funding from a variety of sponsors, usually the same sponsors that would likely provide satellite resources. This reliance on funding from the same sources that would provide satellites could perhaps create a potential conflict of interests. As the pooling system would likely be used in near real-time to help save lives and mitigate property damage, it is unclear if such a potentially important mission would be turned over to an organization created as a non-profit, instead of an organization that had more oversight. There is precedent for non-profits to be responsible for missions that save lives, with one such example being the Red Cross. However, organizations such as the Red Cross are usually created through private movements.

Private Companies

Private companies or coalitions of private companies have long played a large and important role in Earth observation from space. Private companies have filled every role, supporting government agencies in designing and building satellites; designing, building and marketing privately owned satellites; and managing and operating government space operations. The commercial aerospace sector is large and staffed with high quality personnel capable of creating and implementing the technical capabilities necessary for a pooling system. There is also plenty of operational and managerial experience in the private sector to operate a pooling system. Companies without satellites resources that would be useful to the pooling system could also possibly design an effective and unbiased pooling system.

However, as the pooling system would be used to save lives and mitigate property damage, there may be little inclination to turning operation of a pooling system completely over to a for-profit company. Additionally, as many of the satellite resources that would be used in the pooling system are publicly owned, the exclusive use of these resources by a company operating them for-profit would go against previous precedent and law governing the use of public property. The attractiveness from the commercial side is also questionable. A pooling system would be opening a new and untried market. Consumers in the market are used to getting weather related information or research data for either no cost or at near marginal costs. Limiting the access to the data provided would defeat the purpose of the pooling system and may not even be operationally practical. Without limiting access to the information, it could be difficult to encourage a private company to take on the risk of developing and operating a pooling system.

Public Private Partnerships

Public-private partnerships (PPP) are a relatively new organizational tool in the area of satellite Earth observation. Previous domestic examples include PPP for the operation and data dissemination for Landsat 7 with Earth Observing Satellite Company and SeaWIFS design and deployment with Orbital Science, Inc. The purpose of PPPs are to join government and private industry together in an organization that leverages the strengths of both institutions. Government bears much of the risk and supplies a market, while industry provides flexibility and superior managerial skills. Combining public and private institutions also increases the capital base that is available for any program over what would be available if either government or industry pursued a program by themselves. An increased capital base increases the chances that a program will be pursued, thereby increasing the chances that particular information and abilities will be provided for. Another strength of PPP is the flexibility provided in allowing the commercialization of services and products that would otherwise not be provided. Properly designed, a PPP can allow a private company to efficiently supply operational and managerial skills in a program, while commercializing additional services for profit that otherwise would not be provided if government was solely operating the program. PPPs also increase the talent base that can be drawn upon to create and operate a pooling system.

While PPP have several strengths, there are also many difficulties in successful implementation. First, public agencies must have the statutory authority to participate in a PPP that may hand over responsibility or resources to a private firm [28]. Coordination between private companies and public agencies is also a non-trivial matter, as each institution has different missions and different cultures. This coordination is especially difficult if consumers are a third stakeholder in the system, such as research scientists interested in data or the general public's need for information that can be effectively used in near real-time. It is also unclear whether a pooling system or the information that could be obtained from one could form a commercially viable market, without which PPP are usually doomed to failure [28].

8.4 Evaluation of Organizational Models for a Pooling System

The above organizational models were evaluated to determine which would work well as a basis for the pooling organization. A simple evaluation consisting of four metrics was chosen. The metrics chosen were; ability to technically create pooling system, ability to implement pooling system, ability to operate pooling system and cost effectiveness of pooling system operations. Each organizational model was evaluated qualitatively on these four metrics with either a good (+), poor (-) or neutral (0) score assigned to each. A summary table of the evaluation is proved, along with a brief explanation.

Metrics

Below is a discussion of each of the metrics used to evaluate the organizational models discussed above.

Technical Ability to Create Pooling System

Several technical hurdles exist in the design of the pooling system. At a minimum, substantial technical hurdles exist in designing a system that is capable of scheduling and planning satellite operations for a widely disparate set of satellites, taking into account the constraints imposed by individual primary organizations. Additionally, the ability to determine what Earth based phenomenon should be observed, the priority of observations, collection of the necessary information and processing and dissemination of the information into a form that is useful and delivered in a timely manner are all substantial technical problems that must be addressed.

Ability to Implement Pooling System

The ability to coordinate primary organizations in a manner that secures their satellite resources for use in the poling system is considered the largest institutional challenge. It is anticipated that primary organizations will be reluctant to donate their satellite resources to be placed under control of another organization, even for a temporary time. The ability to work out an arrangement that secures enough satellite resources for use in the pooling system is critical to the system's success.

Ability to Operate Pooling System

Operation of the pooling system will be a constant task. Closely related to the operation of the pooling system is being responsive to the appearance of new Earth based phenomenon that warrant observation and the ability to provide the large and differing amounts of data collected in a manner that is useful and timely. It is anticipated that the pooling organization may not be responsible for the processing and dissemination of information to all stakeholders, but would need to interface with the organization that does fulfill this function.

Cost Effectiveness

The ability to operate the pooling system for a reasonable cost is always a priority. Included in the metric of cost effectiveness is the ability of the pooling organization to engage in additional activities that increases the investment of the pooling system. Commercial activities are one potential way in which additional benefits could be realized from the pooling system.

Table 8.1. Evaluation of organizational models for pooling organization.

Organizational Model	Tech. Ability	Imp. Ability	Opert'l Ability	Cost Eff.
Existing Government Research Agency	+	-	-	-
Existing Government Operational Agency	0	-	+	0
New Government Coalition	+	0	+	-
New Government Agency	0	-	+	0
Academic Institution	+	-	0	0
Non-For-Profit Institution	0	-	0	0
Private Sector	+	0	+	+
Public-Private Partnership	+	+	+	+

Overview of Evaluation

An overview of how organizational models were evaluated along each of the four metrics is discussed below. Organizational models are not discussed one by one, but rather general results are discussed.

Technical Ability to Create Pooling System

It is believed that while there are substantial technical problems associated with creating a pooling system, that any of the organizational models could likely assemble the personnel with the talent to solve these problems. Some organizational models may be slightly better than others, but technical problems were not seen as a "show stopper" for the use of any of these organizational models.

Ability to Implement Pooling System

The ability to implement a pooling system is viewed as the largest barrier to making a pooling organization possible. Few organizational models fared well in this metric. Current government organizations were ranked poorly as political infighting between government agencies seen as trying to take other agencies' resources caused these organizational models to rank poorly in this category. Academic and non-profit institutions were seen as being too small and not politically connected enough to overcome resistance that will likely be offered from existing organizations. Private companies formed or contracted for the purpose of creating and implementing a pooling system would likely face the same type of political resistance that existing government agencies would face. The best organizational model for implementing a pooling system as seen as an organization based on a PPP. This is because the PPP would bring in an organization that was non-biased towards any of the existing primary organizations, but would still have ties to public agencies, ideally actively including them in the implementation decision making process.

Ability to Operate Pooling System

Operation of the pooling system was seen as a task that most of the organizational models could handle reasonable well. The only organizational model that was ranked unfavorably in this metric was research agencies, as extended operations are usually outside their mission focus.

Cost Effectiveness

Private sector and PPP organizational models were seen as the most cost effective for pooling system operations. This is primary for the reason that private sector managerial experience is usually considered more cost effective, due to the flexibility afforded to the private sector. Additionally, the ability for private companies to identify and exploit different commercial opportunities while operating the pooling system helps increase system value, thereby increasing cost effectiveness.

8.5 Selection of Organizational Model

As PPPs were viewed to be the most effective organizational model studied for pooling systems, a PPP is recommended for additional study. The following sections examine PPP in more detail.

Overview of Public-Private Partnerships

There are a range of public-private partnership arrangements that could be implemented to guide the pooling organization [26]. These span a range from heavy weighted towards the public sector to heavily weighed towards the private sector. The following section will briefly describe several common PPP types that are applicable to the pooling organization. These are summarized below in Table 8.2.

Table 8.2. Types of PPP arrangements and the degree of privatization.

Public Private Partnership Arrangement	Degree of Privatization
Public Authority	Public Sector Oriented
Contract Based	
Lease Build Operate (LBO)	

Build Transfer Operate (BTO)	
Build Operate Transfer (BOT)	
Buy Build Operate (BBO)	
Build Own Operate (BOO)	Private Sector Oriented

Public Authority

Usually a public organization that is created to act more like a private sector organization. The desire is to create an organization with improved efficiency than what is traditionally found in the public sector. This is usually done with professional mid-level managers, emphasis on relating to target audience as customers, and covering costs of doing business through application of tariffs. Public authorities are often found running large public infrastructure programs, such as transportation, water or electricity.

Contract Based

Contracts are used to delegate services or operations out to private firms. At their core, all contractual arrangements keeps the fiscal risk and ultimate authority with the public agency supplying the contract. The purpose of the contract is to outsource various aspects of the agencies' activities out to the private sector to take advantage of superior or more cost effective services. Services provided can span from support, like janitorial or ticketing, to more essential services where the private firm is contracted to handle all day to day operations and maintenance (called an O & M contract).

Lease Build Operate (LBO)

Long term leases are given to public firms to operate and manage publicly financed and/or built facilities. The absolute ownership of any facility is kept in the public domain, but all responsibility for operations and management are turned over to the private company. Usually the contract is for the very long-term, to reduce uncertainty about future operations and to encourage the private firm to invest in the program.

Build Transfer Operate (BTO)

Similar to the LBO, except the private firm builds the facility instead of taking control of a pre-existing facility. The facility can be funded by the private firm, which after building it transfers ownership to the public entity. The private firm then receives a long term contract for operation. The contract is again very long-term, to encourage private investment.

Build Operate Transfer (BOT)

Similar to the BTO arrangement except that the private firm retains ownership. The private firm finances, builds and operates the program, deriving its revenue from operations directly from the customer base. The private firm is granted the right to pursue this venture from the public firm, which limits competition, making the venture worthwhile to the private firm. After some time, the facility is turned over to the public sector. Because ownership is kept with the private company for so long, this PPP arrangement opens the private company up to legal liabilities, whereas in the above arrangements they are more insulated from liabilities as ownership resides in the public sector. A benefit to the public sector is that by granting the right to engage in commercial activities, the public sector maintains a high degree of control over how the market is exploited.

Buy Build Operate (BBO)

Identical to the BOT arrangement except that instead of building a facility, the private interest buys an existing public capability permanently. Again, revenues are derived directly from customers and the public sector maintains a degree of control by limiting entrants into the market.

Build Own Operate (BOO)

This is the closest arrangement to a purely private venture. A private firm is granted the right to develop and operate their business in perpetuity from the public sector. An example of this would be a toll road.

Public-Private Partnership for the Pooling Organization

While many types of PPP's are available, a Build-Transfer-Operate (BTO) arrangement has been selected as the partnership type most applicable to the needs of creating and operating the pooling organization. The BTO partnership will join both government agencies and at least one commercial sector organization together to meet the needs for the pooling organization. On the government side, several agencies in the U.S. federal government are expected to play a major role, along with international government agencies, in working with commercial sector partners.

Below is a discussion of the goals that the PPP for the pooling organization are desired to meet. Next is an explanation of why a BTO arrangement was selected as the PPP model for the pooling organizations. Directly following this explanation is a discussion of the design of the BTO partnership.

Pooling Organization Public-Private Partnership Goals

The overarching goal of the pooling organization is to provide a means for better predicting and responding to Earth based phenomenon – especially natural and man-made disasters. As there is currently very little, or no, capability for observing Earth based phenomenon in near real time, there is a long-term goal of creating a self-sustaining capability for responding to this need. The pooling organization PPP arrangement is expected to "lay the ground work" for meeting this need.

One of the driving goals for the PPP is to involve the commercial sector in the development of the pooling organization. While a public sector involvement is necessary to provide for the collective needs of society in responding to Earth based phenomenon, involvement of the commercial sector is necessary to fully realize the benefits being provided by the pooling organization. Private sector involvement is necessary to help develop new products and services, add value to the products and services that are provided by the private sector and meet niche market needs not meet by the public sector.

Examples of value added to public products and services by the private area about in the closely related field of weather prediction. While most weather related information is collected through public sources, the private sector utilizes this information to pass it on to the public (weather forecasts on commercial TV, radio, etc.) and create new products (weather forecasts via internet, mobile phone, etc.). The early participation of the commercial sector through the PPP has the goal of speeding up the process of increasing the value of information provided by the pooling organization. The desired result is that the concept of near real time observation of Earth based phenomenon would be more quickly adopted, utilized and embraced as the envisioned host of new products and services coming from this observation would become essential.

Another goal of the PPP is to increase the probability of designing a good pooling organization capability. As the pooling organization will be dependent on satellite resources from a variety of organizations, most likely the majority of which are government organizations, there is the need to bring in a "neutral third party" to help design and operate the pooling organization. The use of a commercial sector organization in this role has the goal of reducing the impact of the infighting that could occur between government organizations, if each were trying to increase their existing organizational mission scope with the new responsibilities of operating the pooling system.

The two goals of faster market penetration and increased probability of a good pooling organization design, along with the generic reasons for engaging in PPP, as discussed above, drive the choice of PPP arrangements and the PPP design, both of which are discussed below.

Rationale for Selecting BTO as Pooling Organization PPP Arrangement

After studying the various types of Public Private Partnerships that are possible, a Build-Transfer-Operate partnership model was selected as an appropriate model. The design of the BTO for the pooling organization is discussed in more detail below. In brief, a commercial organization will secure a competitive contract to build the infrastructure and relationships necessary for the pooling organization, utilizing satellite resources from various organizations. The required infrastructure is then transferred to government ownership and then competitively contracted out to the private sector for operation of the pooling system.

The BTO partnership was selected because it is believed that it joins the public and private sectors together in a manner consistent with fulfilling the goals discussed above. The rationale behind each step in the BTO partnership is explained below.

Build

The use of an independent organization is deemed necessary to build a pooling system that will effectively execute the intended task. Independence is required so that the objectives of the pooling system are not subordinated to another, more established, organization's objectives. This independent organization could have been a private company, a new government agency or a not-for-profit entity. A choice of a new government agency was discarded to avoid the same infighting between government agencies that occurred after the Eisenhower Administration's decision to create NASA (then a new government agency) and transfer resources from more established agencies (like Department of Defense) into it. A not-for-profit organization was discarded for the opposite reason that a private company was decided upon. Which was that the involvement of the private sector could more fully exploit the full potential of data collected from the pooling system.

Transfer

The most expensive parts of fielding a large satellite system for observing Earth based phenomenon are the steps necessary for collecting the data and the creation of new products that use the data collected [24]. While the costs of creating the infrastructure and software associated with the pooling system may not be prohibitive in a fiscal sense, it is believed that creating the organizational linkages between different primary organizations is likely to be hugely expensive. It is believed that transferring ownership of the pooling system to the government is essential for two reasons. First, without direct government involvement and mandate, it is unlikely that any government agencies would participate in the pooling system. Second, the complex organizational interfaces in the pooling organizations present a high risk that the pooling system will not work. It is unlikely that a private company would embark on such a risky project without some involvement by government. As discussed above, there is precedent for government to bear the initial risk for private industry when developing large-scale infrastructure.

Operate

A private company was desired to operate the pooling system to avoid the appearance of infighting between government agencies, to provide superior managerial skills to the pooling systems operation and to be in a position to identify and exploit new commercial opportunities for the data provided from the pooling system. As the pooling system makes use of existing satellites, there is a relatively low capital cost associated with procuring the capabilities for the entire pooling system. Rather, the majority of the costs are anticipated to come from the continuous operation of the pooling system. Here, costs can be kept low by efficient management of operations. Additionally, if the pooling system is not managed well or is not operated efficiently, the commercial partner operating the pooling system can be replaced through a competitive process.

Design of Public-Private Partnership for Pooling System

The PPP envisioned for the pooling system consists of two parts. First is a partnership that has the objective of designing and implementing the pooling system. Second, once the pooling system has been created, a partnership for the purposes of operating the satellite pool should be put into place. The same or different private partners could be chosen for each of these partnerships, though it is envisioned that the firm that designs the pooling system will also be awarded the contract to operate it as well.

Designing and Implementing the Pooling System

A desire to provide the ability to view Earth based phenomenon in near real time is a capability that government must recognize as desirable before any partnership can be formed. The recognition that a PPP may be more desirable than providing for this capability through purely public means is also a necessity. Assuming that these both exist, government must approach the private sector for forming a partnership. On the government side, an organization that does not have a vested interest in the outcome, one that is "above the fray", should be used as the government liaison to industry. Some potential agencies could be the National Science Foundation, the Office of Science and Technology Policy or the US Geological Survey.[8] On the industry side, a request for proposals should be submitted to determine interest from a variety of corporations. It is assumed that one or more corporations would respond. It is also possible that industry partnerships could form and that these g joint ventures may represent industry. One such industry partnership that is currently working with government in a related area is the United Space Alliance, effectively a partnership between Boeing and Lockheed Martin, which manages space shuttle operations.

Once a partnership is formed, the mandate of the first phase of the partnership should be the design of an effective pooling system. It is anticipated that most of the satellites that will comprise the pooling system will be contributed by government agencies. This is due to the fact that the majority of satellites that are currently in orbit for the purposes of

[8] Each of these agencies has strengths and weaknesses associated with using them as a point agency. For example, the National Science Foundation has much experience dealing with the highly technical nature associated with this problem. NSF also has experience distributing funds on its own and would not be beholden to any of the government agencies that have satellite resources. However, NSF is a purely research organization and has little or no experience managing or acting as a partner in a venture that will be very operationally intensive. Similar arguments could be made for the Office of Science and Technology. While OSTP is "above the fray", as it looks at science and technology goals for the entire nation, it has little experience in actually managing an operational system. An agency that does have considerable operational experience with technology systems that are used for a similar purpose is the US Geological Survey. The USGS is an operational agency, though it also has a strong research component. Additionally, the USGS has experience as a consumer in the area of Earth based phenomenon information. However, while the USGS has experience as an operational agency and as a consumer of similar types of data, it has no experience in actually managing satellite systems, which are very technologically complex. Further, as the USGS is a relatively small agency, especially when compared to NASA, the Department of Defense, and NOAA, and because satellite operations is outside the USGS mandate, naming the USGS the lead government partner is likely to meet with resistance from these other agencies with satellite resources.

Earth observation are government owned.[9] Some notable exception of commercially owned and operated Earth imaging satellites are Space Imagining's IKONOS I, DigitalGlobe's QuickBird, Radarsat International's Radarsat 1, Spot Image Company's SPOT 1 and 2, and Earth Observation Satellite Company's Landsat.[10]

One of the first objectives that should be determined is what government agencies would be prime candidates for participating in a pooling system. It is anticipated that agencies that have an agency mission related to near real time observation of Earth based phenomenon would be ideal choices. In this category, the National Oceanic and Atmospheric Administration (NOAA) would be an ideal candidate. The Department of Defense also has a strong interest in weather observations and has historically maintained a fleet of weather observation satellites. Other prime organizations would be one with a research focus. In this category, NASA would be an ideal source of satellites, as would academic satellites operated by universities.[11] Government agencies with considerable satellite resources that very likely would not be interested in participating in any pooling system would be intelligence agencies, such as the Central Intelligence Agency, Defense Intelligence Agency and National Reconnaissance Office.

---

[9] While there are a plethora of commercial satellites, most commercial satellites are communications satellites, which are not directly useful for Earth observations. However, as discussed in the technical portion of this thesis, one of the capabilities that is desirable is for the satellite system to respond to system wide failures in an autonomous and distributed manner. This would require some communications between satellites in the pool. While there would likely be too few satellites tasked to view any one target, some means of relaying information between satellites may be necessary. One method of achieving this could be through use of communication satellites. In this case, as most communication satellites are commercially operated, more commercial involvement in the pooling system may be required.

[10] Of these satellites, only IKONOS I and Quickbird are purely commercial satellites. The other three are part of a public-private partnership in Canada, France and the United States, respectively.

[11] For the three government agencies listed – NOAA, DOD and NASA – participation in a pooling organization may have become easier with recent events. As of 1994, as directed by the National Performance Review and Presidential Decision Directive, NOAA, DOD and NASA weather satellites have been combined into a new, integrated system called the National Polar Orbiting Operational Environmental Satellite System. This system is managed by the tri-agency Integrated Program Office. [sat 6] This convergence of substantial satellite resources under management of one organization is anticipated to make a pooling system more effective, as more satellite resources can be obtained with interactions with fewer organizations.

Also of importance is cooperation with international governments. While most satellites in orbit are US satellites [41], a substantial set of satellite resources should not be overlooked. Cooperation with international partners is consistent with US National Space Policy [34] and precedent[12] [23,39,8,42].

After an initial assessment of what organizations would be likely candidates for participation in a pooling system, the technological challenge of planning and scheduling must be dealt with. The limits of this technology will likely determine the possible usage of satellites. It is anticipated that organizations that may participate will only do so if the amount of time that their satellites are used in the pooling system does not seriously degrade their own agency's primary mission. A planning and scheduling system for use in the pooling organization needs to be developed that will integrate information from each primary organization's satellite availability and effectively use this compilation of information to determine what satellites can be used in the pooling system and when. Without this technical capability, it seems that any organizational links required to form the pooling system will be a "moot point".

Lastly, the initial PPP should address issues related to the provision of what types of information to what customers. As it is likely that most or all of the satellites involved in the pooling system will be government owned and that the operating contract will be paid with government funding, the information provided must be in the public domain. However, it is also desirable that some means of commercializing some aspects of this information is provided. Commercialization opportunities will attract additional private investment and will more fully utilize the public investment, providing an increase in value to the public. It is anticipated that many of the early identified commercial opportunities will be similar to those that are in existence currently. For example, most people get information about the weather through private sources, like commercial TV,

[12] Some notable examples of international cooperation include; The first satellite communications system, IntelSat, which was an international partnership for almost 30 years before being privatized in 2001, and International partnering in the World Data Center System, which was set up after the International Geophysical Year in 1957 to allow countries to share weather information. This partnership was entered into by both the US and USSR and was one of the few partnerships maintained throughout the Cold War.

radio and websites. Additional opportunities will also most likely be supplied as they are currently – under "sweat-of-the-brow", or value added, relationships, which allows commercialization of publicly available information [24,23]. As discussed in the previous chapter, this type of relationship may be problematic if international partners are involved in the pooling system.

Operation of the Pooling System

Once the pooling system has been designed and implemented, including technical organizational aspects, the pooling system will need to be operated by an organization. A PPP is proposed between the government and a private company for operation and maintenance of the pooling system. It is envisioned that the same company that designed and implemented the pooling system will be the same company to operated and maintain the system, but that is not required. Three major activities necessary to engaging in the operation and maintenance of the pooling system have been identified. These include an agreement for use of satellite resources, work on maintaining and building new organizational linkages, and distributing information gathered with the pooling system.

Once the pooling system is in operation, the pool will "borrow" satellites on a temporary basis from organizations for use in the pool. While these satellites will only be used for a small fraction of their design life, some measure of reimbursement must be made between the pooling organization and the primary organizations. It is assumed that the pooling system will have the technology to avoid using satellites when they are being used for a critical aspect of the primary organization's mission. It is recommended that a vending, or lease, agreement be made between each of the primary organizations and the pooling organization. The leases will be short-term leases, for which the pooling organization will compensate the primary organizations for the use of their satellites. The exact amount of compensation is difficult to determine, as the value of the information that otherwise would have been collected is difficult to measure. However, a rough estimate of a compensation floor can be made based on the average life cycle cost of an Earth observing satellite, the life span of the satellite and the amount of time that the pooling organization will use the satellite. This fraction, and sample numbers are shown below.

$$\frac{Satellite\ Life - Cycle\ Cost}{Expected\ Satellite\ Life} * Time\ Satellite\ in\ Pool = Base\ Compensation\ Cost \tag{8.1}$$

$$\frac{\$350M}{10\ years} \cong \$96k\ /\ day$$

Satellites used in the pool will likely only be used from one to seven days, depending on the phenomenon of interest. This places the base compensation cost for each satellite between \$96k and \$672k. As it was found in the technical results of this thesis that groups consisting of six satellites could obtain good coverage of a target, a group would cost between \$576k and \$4M. To place these numbers in context, it has been estimated that an increased accuracy and reliability of information concerning where a hurricane will make landfall would save \$1M for each mile of coastline that did not have to be evacuated.

The short-term use of satellites in this manner accrues relatively small costs that would have to be returned to each of the primary organizations, especially compared to the cost savings in lives saved and economic damage averted. This base cost is only for direct compensation to the primary organizations. Other costs, such as satellite refueling and personnel costs will also be part of the pooling costs, but will not be part of the direct compensation required by the primary organizations.

A major part of the operation of the pooling system will be effort on maintaining the organizational linkages between the pooling organization and the primary organizations. As the pooling system is dependent on the satellites supplied from the primary organizations, maintaining good relationships with primary organizations involved in the pooling system and creating new links with organizations that could contribute to the pooling system is essential. It is anticipated that most new organizations that may be approached for joining the pooling system would be commercial organizations, as opposed to government agencies. This is because the commercial market for satellite imagery of the Earth is rapidly expanding, with several new firms planning on entering this market and up to 12 new Earth imagining satellites slated for launch over the next five years. [28]

As new organizations are brought into the pooling system, the burden on any one primary organization should be reduced. Care must be taken not to marginalize organizations, even if the pooling system is less dependent on them for obtaining satellites. The pooling system may also change as more commercial satellites are brought into the system. An important area of potential change could be the use of data obtained with the pooling system. While data obtained with public sources is provided at a marginal cost to users [24] and is available to all users with no intellectual property protections [24,7,28], commercial satellites do not operate in this manner. While the cost for compensation may be more easily obtained, as it is revenues that would be lost by the commercial primary organization, the lost cost in IP value may be more difficult to determine, and may change the agreements necessary to keep the pooling system operational.

Related to this concern is the design of how information is disseminated once gathered by the pooling system. The flow of information to and from the pooling system is critical in making the pooling system a success. Information must be obtained and processed so that a decision can be quickly made as to what Earth based phenomenon will be observed in a timely manner. Once the phenomenon is observed, the information that is gathered must be processed into a form that is suitable for use and then quickly disseminated so that it can be acted upon. This is especially important if the information is to be used in making decisions that have a short time horizon, such as avoiding tornados or reacting to forest fires.

It is anticipated that one or more additional organizations will be needed to effectively process and disseminate this information. The first type of information that will be obtained is general information related to improved weather forecasts, time-critical information on natural disasters, etc. The second type of information will be less time sensitive information. It is anticipated that a primary consumer of this type of information will be commercial interests that can repackage and add value to the information, allowing it to be used to create new products and sold on the market for profit.

8.6 Discussion of Pooling System Public-Private Partnership Design

Advantages, disadvantages and additional concerns for pooling system PPP are identified and discussed below.

Advantages

Advantages in several areas are possible by applying a PPP to the pooling organization structure. These advantages fall in the areas of; economic, technological, social, and political. Each is discussed below.

Economic

One of the primary benefits for using a PPP is the economic benefits that are possible. The inclusion of both private and public organizations increases the capital base that is available to develop the pooling system and allows each institution to contribute in areas that are its forte, lowering overall system costs. The inclusion of private organizations eases government budgetary constraints by reducing the costs necessary to run a pooling organization, if it can be efficiently run by a private company. Efficient management of the pooling system should lead to an increase in the speed of delivery of the data obtained to the proper customer base.

The value for the money invested is also increased by the inclusion of private companies, if allowed to commercialize some of the date returned from the pooling system. Commercialization allows for a long term increase in the benefits that can be obtained from the pooling system.

Technological

The issues of technology transfer, training and innovation are all technological benefits of a PPP. Technology used for the creation of the pooling system, such as planning and scheduling algorithms, and data processing and dissemination software, can be more efficiently commercialized for use in other sectors through private companies associated with the pooling system. Training in operations for the pooling system will benefit both the pooling organization and the primary organizations that must operate their own satellites the reminder of the time. Innovations in scheduling, planning, operations, data management and dissemination can all be shared between organizations and exploited commercially.

Social Benefits

A pooling system that is more efficiently operated will also benefit the general public by meeting people's needs faster and raising living standards. A more efficient pooling system will increase the effectiveness of the system to respond to more phenomenon in a shorter amount of time, mitigating the consequences from deadly and destructive phenomenon.

Political Benefits

The political benefits for using a PPP allow government to allocate responsibility and minimized distorting influences. Government funding can greatly influence the operation of the pooling system without government management, allowing government to maintain a voice in pooling system use, as opposed to abdicating all responsibility to private interests or other third parties. Political distorting influences are also minimized. Short term political effects, such as technical mandates designed to offer short term political benefits or lack of maintenance funding, are reduced by the existence of the additional barrier of a private company operating the pooling system.

Drawbacks

While there are several advantages for using a PPP for the pooling organizational structure, there are also several drawbacks as well. These include; increased difficulty with dealing with intellectual property obtained from the pooling system, the potential for higher costs and barriers of opposition to PPP in general.

IP Problems

In the US, government data is traditionally open for public use at only the marginal cost of obtaining the data. Private provision of data is always protected to reserve the ability to restrict access to the data so that it can be sold at a profit. The inclusion of both public and private partners in handling data returned from the pooling system complicates the handling of IP protection. Additionally, in the US, private companies can make use of government obtained data cheaply for use in producing "sweet of the brow", or value added, products. This is not the case in Europe, where government data is subject to IP protection, making cheap commercialization difficult. This is a problem if public and private partners from the US and Europe will be included in operating a pooling system.

Potential for Higher Costs

The inclusion of the private sector has the potential to exploit superior managerial skills and flexibility to keep costs low, but there is no guarantee that this will occur. If poorly implemented, the PPP adds another organization to the process and can increase total costs for operating the pooling system.

Barriers of Opposition to PPP

Barriers to the formation of PPP exist that can make the formation of a pooling system PPP difficult. Operational barriers are one such source. These barriers come from institutional difficulties in aligning organizations in the public and private sectors to work together in the common interests of the pooling system. Operational barriers can stem from different organizational structures, cultures, missions, histories and way of doing business. Legislative barriers also exist, which inhibit PPPs. Legislative barriers are laws that prevent the commercialization or the commercial exploitation of publicly financed goods or services. Legislative barriers can also be erected by lawmakers desiring to protect government agencies that they feel would be threatened by the use of a PPP for the pooling organization. Barriers of opposition from different sources can also be expected. Any organization that feels its autonomy, power base or mission is compromised or curtailed by joining the pooling system or participating in the PPP will be expected to oppose creation of the PPP.

Other Potential Areas of Concern

Two other potential areas of concern have been identified for the use of a pooling system. These are the excess commercialization of data and the determination of what data should be released.

Excess Commercialization of Pooling System Data

While some commercialization of some data or the creation of new products obtained from pooling system information is desired and should be encouraged to more fully exploit the pooling system, the over commercialization of pooling system data should be avoided. There may be a strong desire for a commercial pooling system operator to transition as much of the pooling system data from a collective good into a toll good, allowing access to the information to be restricted to allow it to be sold for a profit. The primary purpose of the pooling system was to facilitate the creation of autonomous groups of Earth observing satellites to mitigate effects from Earth based phenomenon. Excess commercialization of the data from the pooling system will overly restrict access and limit the ability of the pooling system to complete the original mission.

Determination of Data to be Released

The provision of data is not always helpful to a particular cause. In fact, supplying certain data can be harmful. This was previously discussed for the case of military security during combat operations and military deployments, but it is not limited to these cases. For example, data collected on the environment that can be used to help track and locate endangered species has very useful scientific properties. However, the wide release of this data into the general domain can actually harm endangered species more than help them, as poachers will also have access to this information and use it to more effectively locate their target [23].

Chapter 9

Conclusions and Future Work

It was found from this study that both the technical and policy problems associated with creating groups of autonomous maneuverable satellites are many but can all be overcome. This study finds that it is possible to achieve the technical results of responding to unexpected events in a timely manner without a substantial increase in fuel usage. It was also determined that the liability issues associated with satellite pooling and organizational cultural inertia are the primary barriers inhibiting organizations from participating in a pool, but that these are possible to overcome as there are examples where similar cross organizational relationships have succeeded, but with great effort. This study finds that the critical barriers that must be resolved before creating a group of autonomous maneuverable Earth observing satellites are not technical in nature, but are legal and cultural changes in organizations.

More detailed conclusions from the technical and policy portions of this study are provided below.

It was found from the technical portion of this study that the integrated planner that was developed, based on the previously developed EPOS 1.0 optimal planner and the previously developed ALLIANCE algorithms, was able to regain at least 95% of the observation time lost due to the occurrence of unexpected events when using only the EPOS 1.0 pre-mission optimal planner. Based on the very decentralized architecture of the reaction planner, the integrated planner developed is only beneficial when the satellites are allowed to exhibit autonomy from a central planner. This results in satellites able to make decisions and implement observation and maneuver plans to maximize its own observations, but offers little guidance on how this decentralized decision would interface with an integrated system that has the goal of maximizing system wide results instead of local satellite results. Because of this, the integrated planner may not be a good choice if system wide coordination with a high degree of interdependence is a high priority.

It was found from the policy portion of this study, several economic, political economy, organizational, legal and political barriers exist that act to inhibit the creation of a pooling system. To overcome these barriers, it was determined that a public-private partnership could provide a means to bring together different stakeholders when creating a pooling system. However, the barriers identified will be difficult to overcome and it is highly uncertain that a pooling system is a feasible means of assembling a group of satellites large enough to observe the number of Earth based phenomenon in real time that would be desired.

A complete list of conclusions and future areas of work is presented below.

- Significant observation time of at least 95% can be regained in most instances after the occurrence of an unexpected event when using the integrated planner.
- The integrated planner provides no observation time regained for most satellites when the unexpected event occurs less than six hours away from the end of the mission.
- Benefits and costs associated with satellites were found to fall into three major classifications that corresponded to satellites that were synchronized, non-synchronized, and in plane with the satellite subjected to the unexpected event.
- Observations can be regained using any re-tasked satellite in 12 hours or less from the occurrence of an unexpected event.

- In all instances the integrated planner was able to re-task a satellite, if a satellite was available and there was enough time left in the mission to perform orbital maneuvers.
- Results for individual satellites vary significantly depending on when in the mission the satellite is re-tasked and which satellite it is replacing.
- It is anticipated that the integrated planner could be easily adapted to respond to the appearance of new targets during the course of the mission.
- It is not anticipated that the integrated planner could be easily integrated into a system that generates optimal plans at the system level.
- It is anticipated that future work in the Earth observing satellite problem will attempt to create an optimal group plan, as opposed to a set of optimal plans for individual satellites that is created using EPOS 1.0. If an optimal group plan is achievable, then the reaction planner utilized here may not be well suited to this application, as the reaction planner makes decisions at the individual satellite level as opposed to the group level.
- The integrated planner is especially suited for applications where decentralized planning is more useful, or even necessary, than the Earth based observation problem. The use of an autonomous satellite network around Mars, or any remote location where central control is difficult, is an example of a possible application.
- More work is needed in finding and applying a metric that generates more realistic orbital maneuver costs for the benefits obtained.
- Additional work is needed in determining if the integrated planner can be modified to act in conjunction with a planner that generates an optimal system plan. It is envisioned that the integrated planner, which is decentralized, would act to temporarily repair the plan until the optimal system planner could regenerate a new optimal plan for the entire system.
- While the planning and scheduling problem is non-trivial, the technical problem of identifying appropriate Earth based phenomenon for observation, compiling observations from several different satellites and disseminating the information in a useful and timely manner is very possibly much more difficult to solve.

- The number of satellites needed to form enough groups of dynamic satellites to observe a large set of targets will likely be larger than any one organization has control over.
- Due to the cost of satellite procurement, innovative methods of forming enough satellite groups may be necessary.
- Satellite pooling has many technical and policy barriers.
- Technical barriers to satellite pooling revolve mainly around planning and scheduling a disparate set of satellites using different operational software with different sets of constraints.
- Constraints on use of the satellites stem mainly from organizational constraints, such as availability of the satellite to complete its primary use or use of the satellite for missions other than one that it was procured for.
- Policy barriers to satellite usage can be expected to be raised by many stakeholders, such as primary organizations and primary satellite users, in many different forms.
- The greatest challenge to creating and operating a pooling system was identified as the implementation of the pooling system such that all needed primary organizations would contribute satellite to participate in the pool.
- A public-private partnership was identified as the most likely organizational model capable of creating a satellite pool, as it overcame several of the identified barriers of resistance.
- It is unclear if even the use of a public-private partnership could overcome all the identified.

Given the results of this study, it appears that the integrated planner is capable of accomplishing the goal of re-tasking satellites so that observation time that would otherwise be lost due to the occurrence of unexpected events is regained. Additional work is necessary to further determine if the integrated planner is a desirable choice for use with any actual planning system.

Given the results of this study, it appears that between the technical and policy challenges that exist in creating and operating dynamic groups of Earth observing satellites, the policy challenges will be much more difficult to overcome. Because of the difficulty in overcoming the barriers associated with satellite pooling and because satellites are currently too expensive for one organization to procure enough satellites capable of providing the needed coverage, a different solution may be required. One such possibility would be pushing new satellite technology that would reduce the costs of satellites to the point that one organization could control al the satellites that are needed to complete this mission. One such possibility is in the use of fleets of microsatellites, though the technological issues associated with microsatellites and the political support for microsatellite programs is currently in question.

References

[1] Abramson, Mark, et al. "The Design and Implementation of Draper's Earth Phenomena Observing System (EPOS 1.0)." *Space 2001*, Albuquerque, NM. August 28-30, 2001.

[2] Ancona, Deborah, Thomas Kochan, John Van Maanen, Maureen Scully, and D. Eleanor Westney: *Managing for the Future: Organizational Behavior and Processes* , South-Western College Publishing Co., 1996.

[3] Assessment of the Usefulness and Availability of NASA's Earth and Space Science Mission Data, Space Studies Board, Board on Earth Sciences and Resources, National Academy Press, 2002.

[4] Baker, John, Kevin O'Connell, Ray Williamson, Commercial Observation Satellites: At the Leading Edge of Global Transparency, 2001.

[5] Bearden, David, Small Satellite Costs, Crosslink, The Aerospace Corporation, V2 N1 2000/2001.

[6] Cordes, Joseph, Henry Hertzfeld, Ray Williamson, The Socio-Ecnomic Value of Improved Weather and Climate Information, the Space Policy Insitutie, George Washington University, December 2002.

[7] Fair Weather: Effective Partnerships in Weather and Climate Services, Board on Earth Sciences and Resources, Board on Atmospheric Science and Climate, Computer Science and Telecommunications Board, National Academy Press, 2003.

[8] Feder, Barnaby, Satellite Company is Trying Life on its Own, The New York Times, July 23, 2001.

[9] High-Resolution Earth Observations From Space: What are Today's Issues? Workshop hosted by The Space Policy Institute, George Washington University and The National Remote Sensing and Space Law Center, University of Mississippi School of Law, December 18, 2001.

[10] Issues in the Integration of Research and Operational Satellite Systems for Climate Research, Part I. Science and Design, Space Studies Board, National Academy Press, 2000.

[11] Jasani, Bhupendra, Orbiting Spies – Opportunities and Challenges, Space Policy 18, 2002.

[12] Mileti, Dennis. *Disasters by Design: A Reassessment of Natural Hazards in the United States.* Joseph Henry Press, Washington D.C., 1999.

[13] Myers, Karen and Smith, Stephen. "Issues in the Integration of Planning and Scheduling for Enterprise Control." *DARPA-JFACC Symposium on Advance in Enterprise Control*; San Diego, CA, November 1999.

[14] The "Outer Space Treaty", Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies, United Nations, 1967.

[15] Nelson, Jane, Building Partnerships: Cooperation Between the United Nations System and the private Sector, United Nations Department of Public Information, 2001.

[16] Pace, Scott, David Frelinger, Beth Lachman, Arthur Brooks, Mark Gabriele, The Earth Below: Purchasing Science Data and the Role of Public-Private Partnerships, RAND 2000.

[17] Parker, Lynne. "ALLIANCE: An Architecture for Fault Tolerant, Cooperative Control of Heterogeneous Mobile Robots." *Proceedings of the 1994 IEEE/RSJ/GI International Conference on Intelligent Robots and Systems (IROS '94)*, September 1994: 776-783.

[18] Parker, Lynne. "ALLIANCE: An Architecture for Fault Tolerant Multi-Robot Cooperation." *IEEE Transactions on Robotics and Automation*, 14 (2), 1998.

[19] Parker, Lynne. *Heterogeneous Multi-Robot Cooperation.* PhD Thesis; MIT, February 1994.

[20] Parker, Lynne. "L-ALLIANCE: A Mechanism for Adaptive Action Selection in Heterogeneous Multi-Robot Teams." *ORNL/TM-13000*, October 1995.

[21] Parker, Lynne. "L-ALLIANCE: Task-Oriented Multi-Robot Learning in Behavior Based Systems." *Advanced Robotics*, Special Issue on Selected Papers from IROS '96, 11 (4) 1997: 305-322.

[22] Principles Relating to Remote Sensing of the Earth from Outer Space, United Nations

[23] Proceedings of the Workshop on Promoting Access to Scientific and Technical Data for the Public Interest: An Assessment of Policy Options, Committee for a Study on Promoting Access to Scientific and Technical Data for the Public Interest, National Research Council, National Academy Press, 1999.

[24] Resolving Conflicts Arising from the Privatization of Environmental Data, Commission on Geosciences, Environment and Resources, National Academy Press, 2001.

[25] Sauer, Birgit. *Autonomous Mission Scheduling for Satellite Operations. Masters Thesis*; MIT, September 1997.

[26] Savas, E.S. Privatization and Public=Private Partnerships, Chatham House Publihers, New York, 2000.

[27] Ticker, R.L and Azzolini, J.D. *2000 Survey of Distributed Spacecraft Technologies and Architectures for NASA's Earth Science Enterprise in the 2010-2025 Timeframe.* NASA Technical Memorandum 2000-209964, August 2000.

[28] Toward New Partnerships in Remote Sensing: Government, the Private Sector, and Earth Science Research, Steering Committee on Space Applications and Commercialization, National Research Council, National Academy Press, 2002.

[29] Transforming Remote Sensing Data into Information and Applications, Space Studies Board, National Academy Press, 2001.

[30] Williamson, Ray, Henry Hertzfeld, Joseph Cordes, John Logsdon, The Socioeconomic Benefits of Earth Science and Applications Research: Reducing the Risks and Costs of Natural Disasters in the United States, 2001.

[31] Comparison of Landsat Missions. http://landsat.gsfc.nasa.gov/project/Comparison.html

[32] Deep Space 1 Mission Website. http://nmp.jpl.nasa.gov/ds1/

[33] Extreme Weather Source Book 2001. http://sciencepolicy.colorado.edu/sourcebook/composite.html

[34] Fact Sheet, National Space Policy, The White House National Science and Technology Council, 1996, www.ostp.gov/NSTC/html/fs/fs-5.html

[35] Mission and Spacecraft Library. http://leonardo.jpl.nasa.gov/msl/QuickLooks/seastarQL.html

[36] The National Polar-orbiting Operational Environmental Satellite System, www.publicaffaris.noaa.gov/grounders/npoess.html

[37] NOAA Strategic Plan, pg. 3. http://www.publicaffairs.noaa.gov/98strpln.pdf

[38] NORAD Two-Line Elements for the SeaStar. http://plankton.gsfc.nasa.gov/sel/

[39] Satellite 2002 – Europe, TelecombWeb, 2002, http://www.telecomweb.com/reports/satellite2002-europe/about.htm

[40] SeaWiFS Project - Detailed Description Website. http://seawifs.gsfc.nasa.gov/SEAWIFS/BACKGROUND/SEAWIFS_970_BROCHURE.html.

[41] The United Nations Registry of Space Objects Section A: Main Registry of Satellites and Space Probes, Table 1: Index to registrations http://hea-www.harvard.edu/QEDT/jcm/space/un/un_taba1.html

[42] Whalen, David J. "Communications Satellites: Making the Global Village Possible." Communications Satellites Short History. http://www.hq.nasa.gov/office/pao/history/satcomhistory.html. 1/2/97.

[43] Whelan, David. "Future Space Concepts." Presentation. http://www.darpa.mil/tto/director_briefs/991110_Whelan_OE_Ind_Day.pdf

[44] Used with permission from David Carter, Draper Laboratory.